CRYSTALLOGRAPHIC AND SOLUTION STUDIES ON BOVINE BETA-LACTOGLOBULIN.

by

SUSAN G. HAMBLING

Ph.D.

UNIVERSITY OF EDINBURGH

1990

DECLARATION.

I declare that this thesis was produced by myself, and that the work it describes is my own, except where stated in the text.

QUOTE AND DEDICATION.

When you get caught between the moon
and New York city,
The best that you can do
is fall in love.

{Theme from 'Arthur'}

This thesis is dedicated to all those I love.

ACKNOWLEDGEMENTS.

First and foremost I should like to thank my supervisor, Dr. Lindsay Sawyer, for his help and tireless enthusiasm towards this project.

I would also like to thank Stella Fawcett for her help with the crystallization work; Drs. Sandy Blake, Miroslav Papiz, Steve Rule, Rob Stansfield and Paul Taylor for their assistance with the crystallography and subsequent computing; Dr. Steve Yewdall for access to the lattice X structure and considerable help with the structural comparison; Roman Hlodan for access to, and help with, the CD; and Dr. Hugh McKenzie for many useful comments.

I also wish to thank the Science and Engineering Research Council for financial assistance for 3 years of this project, the Laboratory of the Government Chemist for access to computing facilities with which to prepare this thesis, my friends for their support especially during the writing up (Dave Watson, Michelle Cunningham, Steve Proctor, Glyn Haylor and Michael Porter), and to Dr. Supavadee and Chris Amatayakul-Chantler for their help throughout.

ABSTRACT

CRYSTALLOGRAPHIC AND SOLUTION STUDIES ON BOVINE BETA-LACTOGLOBULIN.

S.G.HAMBLING

β-lactoglobulin(BLG) is the major whey protein in many ruminant and non-ruminant milks, and, although it has been studied extensively, its function in vivo remains a mystery. Several of the properties of this protein which may provide information on its function are: the distribution of BLG amongst species, the structural homology of BLG to serum retinol binding protein and two bilin binding proteins, its inclusion in a superfamily of transport proteins, the nature of the hydrophobic binding sites, and the pH-induced reversible conformational change (the Tanford transition) which bovine BLG is known to undergo over the physiological pH. These are examined in this thesis.

The preliminary crystal structure of bovine BLG lattice Y at pH 7.8 (space group $B22_12$: a=55.7A, b=67.2A, c=81.7A) had been determined previously. This was improved by the inclusion of more data, and refinement by manual model-building using the molecular graphics program FRODO, least-squares, energy minimization and molecular dynamics. The protein electron density map generated at 2.8A resolution was used to examine the positions of the antigenic regions, and the environments of specific amino acids.

A comparison of this structure with the lattice X structure (space group P1: a=37.8A, b=49.6A, c=56.6A, α=123.4^{o}, β=97.3^{o}, γ=103.7^{o}) determined at pH 6.5 [141] was undertaken, so as to offer a molecular explanation for the conformational change which occurs between these pHs in solution. This transition was known to involve an anomalous carboxyl, and the identity of the residue is discussed. The free cysteine, phenylalanine and tyrosine residues were also implicated. Solution studies, using the techniques of polarimetry, circular dichroism, and tryptophan fluorescence, were carried out to confirm the nature of the Tanford transition.

The inclusion of BLG in a superfamily of transport proteins, and its isolation with fatty acid bound, indicated that its function could be that of a small molecule carrier. The binding of retinol, p-nitrophenyl phosphate and biliverdin to BLG was investigated, to elucidate the nature and importance of the two proposed hydrophobic binding sites - the cavity of the β-barrel, and the external channel near the α-helix.

The relationship between ligand binding to BLG and the Tanford transition is discussed, and a function for this protein is proposed.

CONTENTS

LIST OF ABBREVIATIONS

Proteins

BLG : β-lactoglobulin
RBP : Serum retinol binding protein
INCYN : Insecticyanin
BBP : Bilin binding protein
WAP : Whey acidic protein
PP14 : Placental protein 14
PURP : Purpurin
FABP : Fatty acid binding protein
IFABP : Intestinal fatty acid binding protein
CYT-C : Cytochrome C
OBP : Odorant binding protein
APOD : Apolipoprotein-D
CRCYN-A : Crustacyanin chain A
CRCYN-C : Crustacyanin chain C
A1UG : α_1-microglobulin
Protein HC : Protein heterogeneous-in-charge
Frog BG : Frog protein from Bowman glands
A1AG : α_1-acid glycoprotein
ESP : Epididymal secretory protein
A2UG : α_{2u}-globulin
MUP : Major urinary protein
C8G : Complement 8γ.
P2 : P2 myelin protein

Amino acid codes

Gly, G : Glycine
Ala, A : Alanine
Val, V : Valine
Leu, L : Leucine
Ile, I : Isoleucine
Lys, K : Lysine
Arg, R : Arginine
His, H : Histidine
Tyr, Y : Tyrosine

Trp, W : Tryptophan
Phe, F : Phenylalanine
Cys, C : Cysteine
Met, M : Methionine
Ser, S : Serine
Thr, T : Threonine
Asp, D : Aspartic acid
Glu, E : Glutamic acid
Asn, N : Asparagine
Gln, Q : Glutamine
Pro, P : Proline

Crystallography

R_f : Radius of the film for recording diffraction data
(x_f, y_f) : Coordinates of a diffraction spot on film
XF : Crystal-to-film distance
(x*, y*, z*) : Coordinates of a point in reciprocal space
θ : The Bragg angle

hkl : Miller indices of a reflection
d* : Distance of the reflection (hkl) from the origin of reciprocal space
(X_F, Y_F) : Axes of the film cassette
λ : Wavelength
d : Distance between adjacent reflecting planes in the crystal, and is equivalent to resolution
$\Delta\phi$: Oscillation range recorded on any one film
Δ : Reflecting range of the crystal, and equal to the sum of the beam divergence and crystal mosaicity
1* : Reciprocal cell parameter along the direction of the beam
SRS : Synchrotron radiation source
A : Angstroms
CCP4 : Crystallographic computational project 4
(X_D, Y_D) : Axes of the rotating drum used in scanning
LS : Least squares

(Xc, Yc, Zc) : Orthogonal camera coordinate system

[A] : Orientation matrix relating the camera coordinate system to the reciprocal lattice coordinate system

ϕ_x, ϕ_y, ϕ_z : Missetting angles

Rc_i : Calculated distance of reflection i from Ewald sphere

Ro_i : Observed distance of reflection i from Ewald sphere

R_{scale} : Σ (Fa-Fb) /Σ Fa where Fa and Fb are structure factor amplitudes from the scaled datasets.

B_i : Temperature factor

K_i' : Initial scale factor

K_i : Overall scale factor

Pz : Polarization factor

α : Angle between the rotation axis and the vector d*

ODc : Corrected optical density

ODm : Measured optical density

ODmax : Maximum optical density recorded on a single film

I_o : Observed intensity

I_c : Corrected intensity

N : Number of films in a pack

L : Lorentz correction factor

SD : Standard deviation

R_{merge} : Σ (F_1 - F_2)/Σ F_1 where F_1 and F_2 are the structure amplitudes for symmetry-related/repeat reflections

FRCBIAS : Partial bias calculated from (<If>-Ip) / <If>

<If> : Mean intensity of the full reflections

Ip : Mean intensity of the partial reflections

F : Structure factor

F : Amplitude of the structure factor

sigF : Standard deviation of the structure factor

SD(I), sigI : Standard deviation of the intensities

sigobs : Standard deviation of the observed data

siginp : Standard deviation of the scaled data

rms : Root mean square

MMB+GC : Manual model-building with geometrical constraints

MD : Molecular dynamics

I : Intensity of a reflection

$\rho(x,y,z)$: Electron density at position x,y,z

V : Volume of the unit cell

$\alpha(hkl)$: Phase of reflection (hkl)

MIR : Multiple isomorphous replacement

$\mathbf{F}_P$: Structure factor for the protein

F_{PH} : Structure factor amplitude for heavy atom derivative

F_P : Structure factor amplitude for the protein

$P(\alpha)$: The probability of the phase being α.

m : Figure of merit

R_F : R-factor from $\Sigma (F_{obs} - F_{calc}) / \Sigma F_{obs}$

F_{obs} : Observed structure factor

F_{calc} : Calculated structure factor

Solution studies

$S_{20,w}$: Sedimentation coefficient in seconds

$[\alpha]_D$: Observed optical rotation using the Na-D line

w/v : weight by volume

HPLC : High performance liquid chromatography

M : molar

Kd : Dissociation constant

Ka : Association constant

ORD : Optical rotatory dispersion

CD : Circular dichroism

ESR : Electron spin resonance

L : Pathlength in dm

c : Protein concentration

$[\alpha]\lambda$: Specific optical rotation at λ nm

$[\alpha_{obs}]$: Observed optical rotation

A_{280} : Absorbance at 280nm

ΔA : Differential absorption between left-handed and right-handed circularly polarized light

[x'] : Mean residue ellipticity in degrees.cm^2.$dmol^{-1}$

(mrw) : Mean residue weight of the amino acids

Ia : Fluorescence in the absence of quencher

Ip : Fluorescence in the presence of quencher

K_Q : Quenching constant

[Q] : Concentration of the quencher

Keq : Equilibrium constant

IEF-IPG : Isoelectric focusing on immobilized pH gradients

SDS-PAGE : Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Chemicals

PNPP : p-Nitrophenyl phosphate

DEAE : Diethylaminoethyl

Hg(II) : Mercury (II) salt

Ag(I) : Silver (I) salt

Cu(II) : Copper (II) salt

PHMB : p-hydroxymercuribenzoate

PCMB : p-chloro-mercuribenzoate

SDS : Sodium dodecyl sulphate

NEM : N-ethylmaleimide

EtOH : Ethanol

MeOH : Methanol

BuOH : Butanol

$POCl_3$: Phosphorus oxychloride

CCl_4 : Carbon tetrachloride

2,6-MANS : N-methyl-2-anilino-6-napthalene sulphonate

FFA : Free fatty acid

2,4,4-TMP : 2,4,4-trimethylpentan-2-ol

IEBA : β-ionylidene ethylbromoacetate

PTN : Platinum tetranitrite

MMA : Monomercurio-acetate

HNBB : 2-Hydroxy-5-nitro-benzylbromide

Miscellaneous

HM : Hemochorial

EN : Endotheliochorial

EP : Epitheliochorial

N-terminus / N-end : amino terminus
C-terminus / C-end : carboxyl terminus
kDa : Kilodaltons
mRNA : messenger ribonucleic acid
cDNA : complementary deoxyribonucleic acid
H-bond : Hydrogen bond
dm : Decimetres
PE : potential energy
secIgA : secretory IgA
PAS : 0.4M phosphate + 2.8M ammonium sulphate buffered solution
rpm : Revolutions per minute
ft : Foot
bpi : Bytes per inch
bp : base pairs

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1:1 INTRODUCTION.

The biological function of a protein is dependent upon its structure. For the majority of proteins their functions are recognized, and then rationalized in terms of any available structural information. However one exception is β-lactoglobulin(BLG). Its function is unknown, so the determination of its structure was undertaken as a possible means of explaining its properties and elucidating its function.

BLG is the major whey protein secreted in the milk of many species, and is both abundant and easy to isolate. Since the initial isolation of bovine BLG, by Palmer in 1934 [1], a large number of solution studies have been carried out on this protein, and they have been well reviewed [2,3]. Despite these studies the biological function of this protein remains a mystery.

The failure of solution studies to identify the true function of BLG has led to speculation that a crystallographic approach to the problem might be more revealing. Bovine BLG has been crystallized in a variety of forms [4], and those obtained at pHs 6.5 and 7.8 have had their low resolution structures determined by x-ray crystallography [5]. An improvement in the resolution of these structures would make it possible to offer a molecular explanation for the Tanford transition, the conformational change which this protein undergoes in solution over the physiological pH range [6]. The importance of this transition for the biological function of BLG is unclear; although it has been suggested that the structurally homologous protein, serum retinol binding protein(RBP), undergoes a conformational change

upon binding to its receptor [7], and specific receptors for BLG are known to exist [8].

The preliminary tertiary structure of BLG has been shown to have a high degree of structural homology to three other proteins: RBP, insecticyanin(INCYN) and bilin binding protein(BBP). A comparison of their amino acid sequences revealed several conserved regions, and these were then identified in a number of other proteins [9]. Most of the members of this 'superfamily' of proteins transport small, insoluble or labile molecules, indicating that this may be the role for BLG in vivo [239]. The binding of various compounds (those present in milk or natural ligands for other members of this superfamily) to both dimeric eg. bovine, and monomeric eg. pig, BLGs was therefore investigated, as a means of probing the nature of the binding site.

Prior to any new experimental work being undertaken the current knowledge of the distribution and properties of BLG, and its relationship to other proteins was reviewed [24].

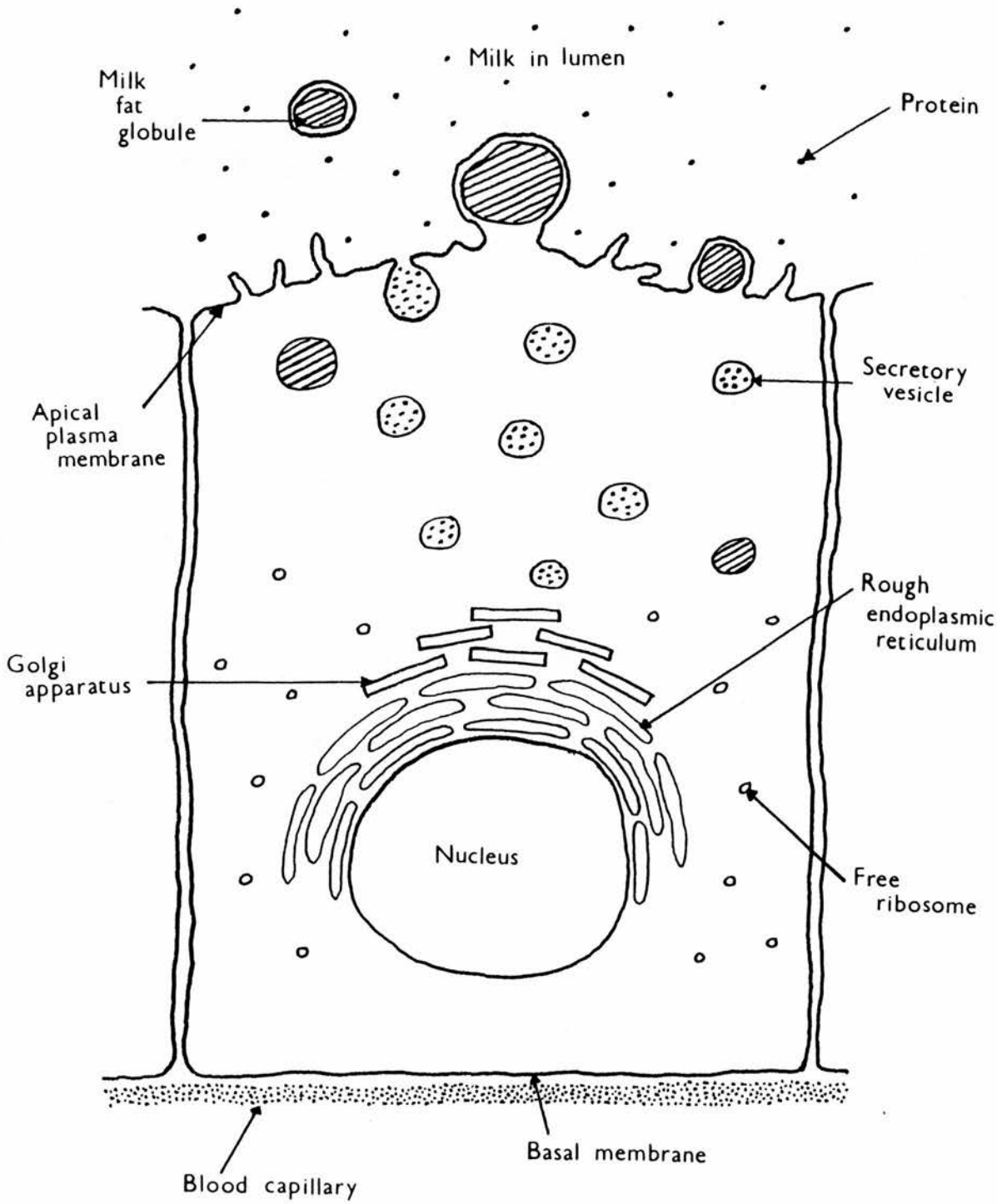
1:2 BIOSYNTHESIS AND SECRETION OF BLG.

Major lactoproteins, including BLG, are biosynthesized within the secretory epithelial cells of the mammary gland {Figure 1:1} [10], under the control of prolactin - a hormone which inhibits the destruction of mRNA coding for these milk proteins. mRNA coding for BLG, which is specific to mammary tissue, is translated at a ribosomal site, to yield a 180 amino acid pre-BLG. Pre-BLGs have been isolated from the mammary glands of cow [11], sheep and pig [12], using cell-free translation systems.

The signal peptides consist of 18 highly conserved amino acids, which are very hydrophobic, possibly α-

Figure 1:1

SIMPLIFIED DIAGRAMMATIC REPRESENTATION OF A SECRETORY EPITHELIAL CELL IN THE LACTATING MAMMARY GLAND.



helical, and have a basic residue near their N-terminus that could form an ionic interaction with charged groups on the surface of a membrane. Their binding to the rough endoplasmic reticulum anchors the ribosome, and enables the pre-BLG to be piloted through the membrane into the cisternal space. A membrane-bound proteinase then cleaves off the signal peptide after the Ala at its C-terminus, to generate the mature 162 amino acid protein.

The mature protein can then undergo post-translational modifications. N-glycosylation is believed to occur in the cisternal space, by the transfer of an oligosaccharide from membrane lipids to an Asn residue in an Asn-X-Ser/Thr sequence, eg. to Asn-28 in BLG from the Australian Droughtmaster breed of cattle [13]. The sequence Asn-Pro-Thr (152-154) occurs in all bovine BLGs but appears to never be glycosylated. This could be explained from the preliminary crystal structure at pH 7.8 [8], which revealed that the Asn is situated between the two subunits and would be sterically prevented from remaining glycosylated upon the association of the subunits in the final assembly. The presence of a disulphide isomerase is likely to ensure correct folding of the protein before its transport to the Golgi apparatus, and incorporation into secretory vesicles. The native protein is then secreted into the lumen, where it accumulates in the milk prior to removal by the young.

The composition of the milk secreted varies according to the species and stage of lactation. In the case of cow, which transfers passive immunity to its young via lacteal secretions, the prepartum lacteal whey contains a high quantity of IgG_1, the level falling before birth as the amount of BLG increases from 2.4 to 5mg/ml. After parturition the levels of both compounds fall, the concentration of BLG returning to its original value within a week [51], before attaining another maximum at mid-lactation.

1:3 DISTRIBUTION OF BLG.

The distribution of BLG amongst the milks of many species has led to some speculation on the role of BLG.

Since the initial preparation of BLG from bovine milk in 1934 [1] dimeric BLGs have been isolated from the milk of a number of other ruminants, and monomeric BLGs have been purified from the milk of several non-ruminant species. BLG has also been detected in the milk of other species, but the state of its association is uncertain. It seems likely that any BLGs detected by antibovine BLG antisera are dimeric, as there is known to be little cross-reactivity between this antisera and monomeric BLGs. Hence the lack of cross-reactivity does not rule out the presence of a monomeric BLG. In two cases, donkey [43] and peccary [17,65], there was no cross-reactivity with antibovine BLG antisera, but upon using antisera raised against monomeric BLGs cross-reactivity was observed. The molecular weight of kangaroo BLG has been quoted as 17-19kDa [14], suggesting a monomeric structure, although a subsequent report declared, without evidence, that it was dimeric despite its low primary sequence homology compared to bovine BLG (32%) [15]. The actual existence of camel BLG is still in some doubt: its absence was stated [16], and it is known that camel milk does not cross-react with antisera to bovine BLG [17]. However, higher titres of antibodies to bovine BLG have been shown to cross-react with both camel milk and colostrum [18]. This may indicate that there is a monomeric BLG present. Cats milk has recently been shown to cross-react with antisera raised against both bovine and equine BLGs, although higher titres of the former were required, indicating that a monomeric BLG could be present. This was confirmed by both gel electrophoresis and N-terminal sequencing [49] {Figure 1:2}.

Figure 1:2

DISTRIBUTION OF BLG IN THE MILKS OF VARIOUS SPECIES.

Species	Cross-reaction with antibovine antisera	State of association
Cow {*Bos taurus*}	+	Di[87]
Oxen {*Bos javanicus*}	+	Di[75]
Yak {*Bos grunniens*}	+	Di[32]
Zebu {*Bos indicus*}	+[17]	?Di*
Waterbuffalo {*Bubalus arnee*}	+	Di [60]
Buffalo {*Bubalis bubalis*}	+[17]	?Di*
Bison {*Bison bison*}	+[17]	?Di*
Musk ox {*Ovibos moschatus*}	+[17]	?Di*
Eland {*Taurotragus oryx*}	+[17]	?Di*
Goat {*Capra hircus*}	+	Di[61]
Sheep {*Ovis aries*}	+	Di[77]
Mouflon {*Ovis ammon musimon*}	+	Di[37]
Red deer {*Cervus elaphus L.*}	+[17]	?Di*
European elk {*Alces alces L.*}	+[17]	?Di*
Reindeer {*Rangifer tarandus L.*}	+[17]	?Di*
White-tailed deer {*Odocoileus virginianus*}	+[17]	?Di*
Fallow deer {*Dama dama*}	+[17]	?Di*
Caribou {*Rangifer arcticus*}	+[17]	?Di*
Giraffe {*Giraffa camelopardalis*}	+[17]	?Di*
Okapi {*Okapia johnstoni*}	+[17]	?Di*
Prong-horn antelope {*Antilocapra americana*}	+[17]	?Di*
Camel {*Camelus dromedarius*}	+/-[18]	?Mono*
Llama {*Lama glama L.*}	-[48]	None[48]

Donkey {*Equinus asinus*}	-[43]*	Mono[43]
Horse {*Equus caballus*}	+/-[83b,42]	Mono[42]
Zebra {*Equus quagga*}	-[17]	?Mono/None*
Rhinoceros {*Diceros bicornis*}	-[17]	?Mono/None*
Peccary {*Pecari tajacu*}	-[17]*	Mono[65]
Pig {*Sus scrofa domestica*}	+/-[83b]	Mono[41]
Dolphin {*Tursiops truncatus*}	nm	Mono[44]
Manatee {*Trichechus manatus latirostris*}	nm	Mono[44]
Beagle {*Canis familiaris*}	nm	Mono[34,44]
Cat {*Felis catus*}	+/-[49]*	Mono[49]
Grey kangaroo {*Macropus giganteus*}	nm	Mono[14]*
Red kangaroo {*Macropus rufus*}	nm	Mono[14]
Northern fur seal {*Callorhinus ursinus*}	nm	Unknown[46]
Human {*Homo sapiens*}	+/-[20a]	None[19]*
Chimpanzee {*Pan troglodytes*}	nm	None[39]
Lemur {*Galago crassicaudatus*}	nm	None[39]
Mouse {*Mus musculus*}	-	None[29]
Rat {*Ratus norvegicus*}	nm	None[26]
Guinea-pig {*Cavia porcellus*}	nm	None[47]
Rabbit {*Oryctolagus cuniculus*}	-[17]	None

Key

+ = BLG detected by antibovine BLG antisera
- = BLG not detected by antibovine BLG antisera
+/- = BLG detected by antibovine BLG antisera, but only at higher titres
nm = cross-reactivity not measured
None = no BLG present
Unknown = BLG present, but state of association unknown
Di = Dimeric BLG detected
Mono = Monomeric BLG detected
* = see text

There is believed to be no BLG in human milk [19] - despite the isolation of a 13.5kDa whey protein with pI 5.1 [20b], and the characterization of a protein from human colostrum with a pI 6.1, a molecular weight around 18kDa and an amino acid composition similar to that of pig BLG [21]. Both these proteins were detected by immunoprecipitation with antibovine BLG antisera, and are most likely to be β_2-microglobulin [33] or fragments of human lactoferrin [22a,22b], which are known to cross-react with this antisera. Both β_2-microglobulin and lactoferrin contain a region which shows homology to the residues 124-140 of bovine BLG, suggesting that some of these amino acids may be part of an antigenic region of BLG. This will be discussed further in Chapter 5:4.

Radioimmunoassays have suggested that low concentrations of bovine BLG, or antigenic peptides of it, (5-800μg/l), are present in human milk, having been transferred from ingested cows milk in the mothers diet across the intestinal barrier, virtually intact, and via the blood stream to the mammary gland [22b,23].

BLG is also believed to be absent from the milk of rodents. Both rat and mouse milks contain a major whey acidic protein(WAP) which has a similar molecular weight to, but is not, a BLG. There is no evidence of hybridization between cDNA from a rat mammary library and cDNA for sheep BLG [25], and mouse WAP shows no primary sequence homology to bovine BLG [26]. Mouse BLG has been reported to be the major lactation protein in murine milk [27], but since it fails to cross-react with bovine BLG antisera [17] it is more likely to be the WAP. A recent publication has confirmed both the abundance of WAP (2g/l) and the absence of BLG [29]. The 'mouse BLG' that has become commercially available could either be the WAP, or BLG from a transgenic strain of mice. The successful expression of sheep BLG in transgenic mice has been reported [30].

A possible correlation between the distribution of species (amongst the orders within the placentalia class of mammals) that produce BLG and their placental type has been proposed [28,141] {Figure 1:3}. (Placental types are defined according to the degree of contact between the uterus and chorion - for more details see Appendix 1). It appears that this correlation may be extended to include the kind of immunoglobulin transferred from mother to offspring {Figure 1:4}. Those species which have a high degree of intimacy between the uterus and chorion during gestation (placental type = hemochorial) are able to transfer IgG to their young in utero, and produce no BLG in their milk; whereas those species with a lower degree of contact between maternal and fetal circulations (placental type = epitheliochorial) cannot transfer immunoglobulins in utero, and thus transfer IgG via colostrum and have BLG in their milk. The relevance of these observations to the function of BLG is unclear.

As BLG is a milk-specific protein its function is likely to be important for the health of mammalian offspring in their early days. It seems reasonable to propose that those species with no BLG in their milk, but a high degree of intimacy between uterus and chorion, might produce another protein with a similar function that could be transferred to the young in utero. Recently a placental protein, placental protein 14(PP14) - also known as α_2-pregnancy associated endometrial globulin, and endometrial protein 15 [53] - has been isolated from human, a species devoid of BLG [54]. This protein, which is synthesized from the mid-luteal phase to the first trimester of pregnancy under the control of progesterone, is very similar to BLG [52]. The coding region of its gene sequence shows 67% homology to that for ovine BLG, whilst its amino acid sequence shows 53% homology to that of horse BLG-I. PP14 may be polymorphic, like some BLGs, as a few differences exist between the primary sequences

Figure 1:3

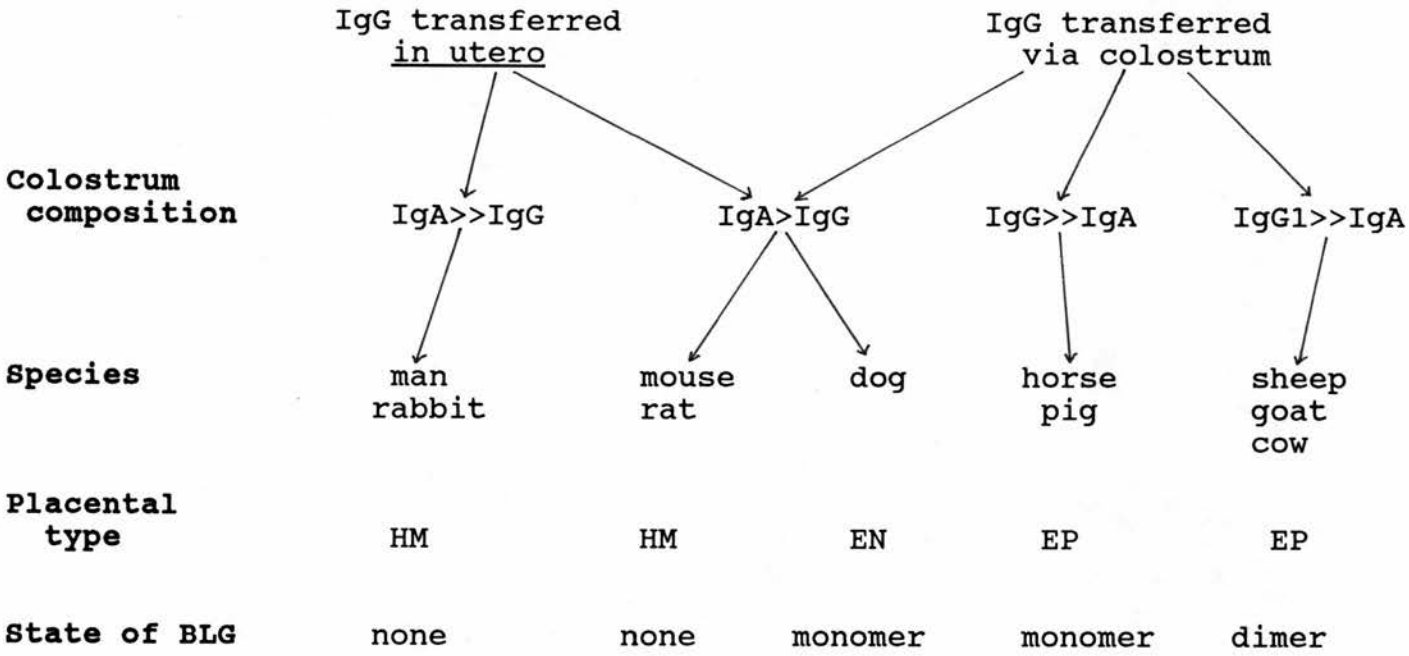
CORRELATION BETWEEN BLG PRODUCTION AND THE PLACENTAL TYPE OF VARIOUS MAMMALS.

Order	Family	Species	BLG present?	Placental type[50]
Primates	Hominidae	man	no	HM
Rodentia	Muridae	rat, mouse	no	HM
	Caviidae	guinea-pig	no	HM
Lagomorpha	Leporidae	rabbit	no	HM
Sirenia	Trichechidae	manatee	M	?HM/EN
Carnivora	Felidae	cat	M	EN
	Canidae	dog	M	EN
Cetacea	Delphinidae	dolphin	M	EP
Perissodactyla	Equidae	horse	M	EP
Artiodactyla	Suidae	pig	M	EP
	Bovidae	cow,sheep	D	EP
	Cervidae	deer	D	EP

Key
no = No BLG present
M = Monomeric BLG present
D = Dimeric BLG present
HM = Hemochorial
EN = Endotheliochorial
EP = Epitheliochorial

Figure 1:4

RELATIONSHIP BETWEEN IMMUNOGLOBULIN TRANSFER, PLACENTAL TYPE, AND THE STATE OF ASSOCIATION OF BLG FROM VARIOUS MAMMALS.



Key
HM = Hemochorial
EN = Endotheliochorial
EP = Epitheliochorial

of the N-terminus region determined by amino acid analysis [55,56,45] and those from several cDNA clones [52]. However all the features thought to be important in BLGs (see Chapter 1:10) are conserved in the complete primary structure of PP14 [52]. It is interesting to note that although unglycosylated PP14 has a similar molecular weight compared to BLGs, it can be glycosylated at three sites, one of which is Asn-28, the single glycosylation site found in BLG-Dr [13].

1:4 ISOLATION OF BLG.

The isolation of BLG from milk is a simple procedure, involving just four stages: the removal of fats, the removal of the caseins, the fractionation of the whey proteins, and the purification of BLG. Since each stage can be carried out in a number of ways, a variety of methods are available for obtaining BLG.

Palmer's original method for the isolation of bovine BLG from skimmed milk [1] was eventually superseded by the method of Aschaffenburg & Drewry [57], which involved an acid precipitation step at pH 2. It was subsequently suggested that the harshness of this step might cause an irreversible conformational change to BLG, and so the method was modified by Armstrong et al. such that the whey proteins were fractionated by acid precipitation at pH 3.5 [58]. A more recent method (Monaco et al. [59]) separated these proteins by DEAE-cellulose chromatography without the pH going below 6.6. Crystals grown from this BLG diffracted to a higher resolution than those grown from BLG purified by the method of Aschaffenburg & Drewry, suggesting that acid precipitation at pH 2 could damage the protein to some extent. Despite this, most commercial preparations of bovine BLG have been subjected to the procedure of Aschaffenburg & Drewry. A summary of

the methods mentioned above is given in Figure 1:5.

Other dimeric BLGs have also been obtained by these methods, or slight variations thereof. Yak BLG [32] has been isolated by the procedure of Aschaffenburg & Drewry, whereas red deer [38], waterbuffalo [60], sheep [36] and goat [61] BLGs have been obtained by modifications of this method. More recently both goat [62] and ovine [63] BLGs have been isolated by a non-salt precipitation method, in which the fat and caseins were removed by centrifugation and then the whey protein mixture was dialysed, lyophilized, and separated by gel filtration.

The isolation of monomeric BLGs from the milk of non-ruminants can be carried out using similar procedures to those described above, with equal ease. Pig BLG has been isolated and purified by three different methods, and these are summarized in Figure 1:6. No comparison appears to have been made on the advantages of the various methods used. The procedure of Kalan et al. [65] appears worst as the many precipitation steps are probably the cause of the low yield observed (25%); whereas that of Kessler and Brew [41] has been successfully used to isolate monomeric BLGs from dolphin, manatee, and beagle milks [44].

The amount of BLG obtained by the various methods depends upon both the procedures used, and the quantity of BLG in the initial milk, which is known to vary with season and time since parturition. The quantities of BLG isolated from the milks of several species are tabulated in Figure 1:7.

1:5 GENETIC VARIANTS OF BLG.

The original fractionation of the whey proteins of bovine milk [1] produced a heterogeneous preparation of BLG. This was due to the presence of two genetic

Figure 1:5

SOME ISOLATION PROCEDURES FOR BOVINE BLGS.

Method	Removal of fat	Removal of caseins	Fractionation of whey proteins	Purification
Palmer [1]	Skim milk used	Acid precipitation	By sodium sulphate precipitation at up to 30°C	Dialysis Recrystallization
Aschaffenburg & Drewry [57]	With sodium sulphate precipitation at 40°C		Acid precipitation at pH 2 to remove aLa and SA, ammonium sulphate precipitation of BLG at pH 6	Dialysis Recrystallization
Armstrong, McKenzie & Sawyer [58]	By ammonium sulphate precipitation at 20°C		Acid precipitation at pH 3.5 to remove aLa and SA, ammonium sulphate precipitation of BLG at pH 6	Dialysis Recrystallization
Monaco et al.[59]	Centrifugation	Calcium chloride precipitation at pH 6.6	Dialysis, DEAE-cellulose chromatography using sodium chloride to elute BLG	Gel filtration Crystallization

Key
aLa = α-lactalbumin
SA = serum albumin

Figure 1:6

ISOLATION PROCEDURES FOR PIG BLGS.

Method	Removal of fat	Removal of caseins	Fractionation of whey proteins	Purification
Kessler & Brew [41]	Acidification		Gel filtration	Ion-exchange chromatography Gel filtration
Kalan, Kraeling & Gerrits [65]	Separator at 40°C	By acid precipitation	By sodium sulphate precipitation to remove aLa, acid precipitation of BLG at pH 3	pH fractionation Ammonium sulphate fractionation Crystallization
Bell, Shaw & McKenzie [66]	Centrifugation	By ammonium sulphate precipitation	Ammonium sulphate precipitation of whey proteins, gel filtration	Ion-exchange chromatography

Key
aLa = α-lactalbumin

Figure 1:7

QUANTITIES OF BLG IN THE MILK OF VARIOUS SPECIES.

Species	Quantity of BLG(mg/ml)	Reference
Bovine	1.8 to 5.0	[1,51]
Goat	1.4	[61]
Sheep	2.8	[29]
Red deer	2.8 to 3.0	[38]
Pig	0.6*	[65]
Dolphin	16.2	[44]
Manatee	14.1	[44]
Beagle	10.1	[44]
Human	0.000005-0.0008**	[23]
Mouse	0	[29]

Comments

* This value appears very low, probably due to the poor fractionation of the porcine whey proteins at pH 3, and the numerous precipitation steps used subsequently.

** It is unlikely that this result refers to human BLG. Radioimmunoassays using antibovine BLG antisera have detected β_2-microglobulin [33], fragments of lactoferrin and bovine BLG in human milk [22b].

variants, A and B, which have been separated [70], sequenced [71], and are the ones most often present in commercial preparations. Other genetic variants of bovine BLG have also been detected, isolated, and sequenced. Cow is currently believed to produce six variants: BLG-A and BLG-B as mentioned above, BLG-C which has only been positively identified in the milk of Jersey cows [72a,72b] but may also be present in South African Nguni cattle [73], BLG-D [74,69], BLG-H [76] and BLG-W which only exhibits a single neutral amino acid substitution compared to BLG-B [67,86]. A cross-breed between cow and zebu, Australian Droughtmaster, was shown to produce BLG variants A, B and a new one BLG-Dr, which is the only BLG known to be glycosylated (it is N-glycosylated at a single site, Asn-28 [13]). Bali cattle (oxen) possess three genetic variants: BLGs E, F and G, the latter containing a single neutral substitution with respect to BLG-E [75]. It has recently been suggested that BLG-E is identical to yak BLG, as position 11 in the latter is now thought to be Asp [75] and not Asn [32]. This implies a close relationship between bovine subgenera.

Other ruminant BLGs also exhibit genetic polymorphism. Red deer possess two BLGs [38], whereas sheep have three codominant alleles coding for the variants BLG-A, BLG-B [77,63,78] and BLG-C [68]. Mouflon BLG is homogeneous, and identical in sequence to ovine BLG-B [37]. Waterbuffalo [60] and goat [62,79] BLGs are both homogeneous by electrophoresis, but the latter (unlike any of the above proteins) contains an amino acid substitution which cannot be due to a single point mutation, namely Asp-130 (GAU/C) in bovine BLG-A becomes Lys-130 (AAA/G) in goat BLG. It was not surprising to discover that this mutation occurred at the beginning of the α-helix in the protein structure, and was situated at an intron/exon boundary in the gene structure [275].

Monomeric BLGs from kangaroo, dolphin, horse and

pig, have also been purified to homogeneity and sequenced. Their sequences diverge considerably from bovine BLG-A and one another, and many of the substitutions observed could not have arisen from single point mutations. Kangaroo BLG is homogeneous [15], whereas dolphin, donkey, horse and pig show genetic polymorphism. Two dolphin BLGs have been detected and both have been partially sequenced [80,44]; whereas one of the two donkey BLGs isolated has been fully sequenced [43]. Equine BLG has been shown to be heterogeneous [81a], despite an earlier report to the contrary [42]. Two variants have been purified and sequenced - the complete primary structure of BLG-I [81b] agrees with its partial sequence determined earlier [42], but reveals only 70% identity to BLG-II [64], suggesting that these variants may be the result of gene duplication. Cat BLG is also heterogeneous - it is currently believed to have up to 12 variants originating from three loci [49].

The heterogeneity of pig BLG was first reported in 1969 [82] and the variants named BLG-A and BLG-B [65]. A later study also isolated two variants: one was believed to be BLG-A, but the other was distinct from BLG-B and named BLG-C [66]. Recently BLG has been purified from sows milk, and studied by two methods [83a]: chromatofocusing yielded three peaks of which only the first, at pH 4.7, was homogeneous, and thus sequenced as pig BLG-I [84]; whereas IEF-IPG gave three bands, BLG-1 BLG-2 and BLG-3 with pIs of 4.4, 4.65 and 4.9 respectively. The amino acid compositions of all these proteins are given {Figure 1:8}, but it is unclear which variants are the same. These discrepancies suggest that many of the pig BLGs mentioned above could be mixtures of more than one variant, implying that phenotyping may be necessary. The existence of a large number of pig BLG variants would not be surprising in view of our knowledge about cat BLGs, and might indicate the presence of gene

duplication. Homology between the first fifty amino acids of BLG-A [66] and BLG-I [84] suggests that these two may be equivalent. The pig BLG to be used in later solution experiments was generously donated to us by C.Holt. It was believed to be swine BLG-B, but this was not fully confirmed by its amino acid composition, as kindly determined for us by A.Cronshaw {Figure 1:8}.

There appears to be no correlation between the species which produce genetic variants of BLG {Figure 1:9}. However those variants from ruminants show greater than 90% sequence homology between one another and other dimeric BLGs {Figure 1:10} and are synthesized under the control of codominant alleles; whereas non-ruminant variants are monomeric, show much lower primary sequence homology to other monomeric and dimeric BLGs (typically 30-70%) {Figure 1:11}, and may originate from gene duplication, as suggested for horse BLGs [64], and detected for cat BLGs [49].

The differences between monomeric and dimeric BLGs will be discussed with respect to their primary sequences, the structure of bovine BLG, and their varying properties in Chapter 5:5.

1:6 SOLUTION STUDIES.

The abundance, ease of preparation, and relative stability of bovine BLG has enabled numerous solution studies to be carried out, using a wide variety of techniques. These experiments have elucidated its physico-chemical behaviour, and provided an insight into the importance of specific amino acids.

1:6:1 AGGREGATION BEHAVIOUR.

The molecular weight of bovine BLG has been measured

Figure 1:8

THE AMINO ACID COMPOSITION OF PIG BLGS.

Variant:	Kalan[65] A	Kalan[65] B	Bell[66] A	Bell[66] C	Conti[84] I	Conti[83a] 1	Conti[83a] 2	Conti[83a] 3	Us*
Amino acid									
Asx	17	17	17	18	17	16	18	18	17
Thr	10	10	10	9	10	10	9	10	10
Ser	10	10	9	9	8	7	9	7	7
Glx	23	23	24	22	24	22	21	24	24
Pro	8	8	8	8	7	7	6	7	8
Gly	3	3	4	4	3	7	5	4	3
Ala	13	14	14	13	13	13	12	13	13
Cys	4	4	4	4	5	4	6	5	nd
Val	13	12	13	13	13	15	10	12	12
Met	4	4	4	4	4	nd	2	3	3
Ile	6	6	6	6	6	6	5	6	6
Leu	24	24	25	24	24	22	27	25	25
Tyr	2	2	2	2	2	nd	5	3	2
Phe	3	3	3	3	3	3	6	3	3
Lys	11	11	11	11	11	14	11	11	12
His	3	3	3	4	3	4	4	4	4
Arg	5	5	5	5	5	5	4	5	6
Trp	1	1	1	1	1	nd	nd	nd	nd

Comments
nd = amino acid residue not detected
* = this amino acid composition was obtained by A.Cronshaw from the pig BLG kindly supplied by C.Holt. This protein was used in later experiments.

Figure 1:9

KNOWN GENETIC VARIANTS OF BLGS.

Species	Number of variants	Variants	Amino acid information
Bovine			
Cow	6	A	ws [71,62]
		B	ws [71,62]
		C	ws [72b]
		D	ws [69]
		H	comp [76]*
		W	ws [86]
Oxen	3	E	ws [75]
		F	ws [75]
		G	ws [75]
Zebu	1	Dr	ws [13]
Yak	1	D_{yak}	ws [32,75]
Sheep	3	A	ws [63]
		B	ws [78]
		C	ws [68]
Red deer	2		none [38]
White-tailed deer	2		none [17]
Horse	2	I	ws [81b]
		II	ws [64]
Donkey	2	I	ws [43]
		II	none [43]
Dolphin	2	1	ps [80]
		2	ps [44]
Pig	?6	A (I)	ws [66,84]
		B	comp [65]
		C	ps [66]
		1	comp [83a]
		2	comp [83a]
		3	comp [83a]
Cat	?12		ps [49]
Dog	>1		ps [44,91]

Key

ws = whole primary sequence known
ps = partial primary sequence known
comp = amino acid composition known but not primary sequence
none = no information on amino acid composition or sequence available
* = the complete primary sequence has been established by A.Conti, S.Ronchi, J.Godovac-Zimmermann et al., but is unpublished [86]

Figure 1:10

COMPARISON OF THE AMINO ACID SEQUENCES OF RUMINANT BLGS.

	Variable amino acid positions																	
BLG	**1**	**11**	**20**	**28**	**45**	**50**	**53**	**56**	**59**	**64**	**78**	**118**	**129**	**130**	**148**	**150**	**158**	**162**
Bovine A	V	D	Y	D	E	P	D	I	Q	D	I	V	D	D	R	S	E	I
Bovine B	V	D	Y	D	E	P	D	I	Q	G	I	A	D	D	R	S	E	I
Bovine C	V	D	Y	D	E	P	D	I	H	G	I	A	D	D	R	S	E	I
Bovine D	V	D	Y	D	Q	P	D	I	Q	G	I	A	D	D	R	S	E	I
Bovine E	V	D	Y	D	E	P	D	I	Q	G	I	A	D	D	R	S	G	I
Bovine F	V	D	Y	D	E	S	D	I	Q	G	I	A	D/Y*		R	S	G	I
Bovine G	V	D	Y	D	E	P	D	I	Q	G	M	A	D	D	R	S	G	I
Bovine Dr	V	D	Y	N*	E	P	D	I	Q	D	I	V	D	D	R	S	E	I
Bovine W	V	D	Y	D	E	P	D	L	Q	G	I	A	D	D	R	S	E	I
Yak	V	N/D*	Y	D	E	P	D	I	Q	G	I	A	D	D	R	S	G	I
Waterbuffalo	I	D	Y	D	E	P	D	I	Q	D	I	A	D	D	R	S	E	V
Goat	I	D	Y	D	E	P	N	I	Q	D	I	A	D	K	R	A	G	V
Sheep A	I	D	Y	D	E	P	N	I	Q	D	I	A	D	N	R	A	G	V
Sheep B	I	D	H	D	E	P	N	I	Q	D	I	A	D	N	R	A	G	V
Sheep C	I	D	Y	D	E	P	N	I	Q	G	I	A	D	N	Q	A	G	V

Key

D/Y* : it is unclear which residue is at which position

N/D* : this residue was believed to be Asn[32], but is now thought to be Asp[75]

N* : represents an Asn residue with a carbohydrate moiety attached. It is within the sequence -NIS- which is typical of a N-glycosylation site.

Figure 1:11

COMPARISON OF THE PRIMARY SEQUENCES OF NON-RUMINANT BLGS AND BOVINE BLG-A.

```
                  *       * *                     **           * *   **   *        *
Bovine A    LIVTQTMKGL DIQKVAGTWY SLAMAASDIS LLDAQSAPLR -VYVEELKPT PEGDLEI-LL QKWENDECAQ
Kangaroo    VENIRSKNDL GVEKFVGSWY LREAAKTMEF -----SIPLF DMDIKEVNLT PEGNLELVLL EKTDR--CVE
Dolphin I   VSVIRTMEDL DIQRVAGTWH SVAMAASDIS LLDTEEAPLR -VNVEELRPT PQGDLEL-FL QK........
Donkey      TNIPQTMQDL DLQEVAGKWH SVAMAASDIS LLDSEEAPLR -VYIEKLRPT PEDNLEI-IL REGENKGCAE
Horse I     TNIPQTMQDL DLQEVAGKWH SVAMAASDIS LLDSEEAPLR -VYIEKLRPT PEDNLEI-IL REGENKGCAE
Horse II    TDIPQTMQDL DLQEVAGRWH SVAMVASDIS LLDSEEAPLR -VYVEELRPT PEGNLEI-IL REGANHACVE
Pig I       VEVTPIMTEL DTQKVAGTWH TVAMAVSDVS LLDAKESPLK -AYVEGLKPT PEGDLEI-LL CKRENDKCAQ

                ** *  *                          *                        *
Bovine A    KKIIAEKTKI PAVFKIDALN ENKVL--VLD TDYKKYLLFC MENSAEPEQS ----LVCQCL VRTPEVDDEA
Kangaroo    KKLLLKKTKK PTEFEIYISS ESSYTFCVME TDYDSYFLFC LYNISDREK- ----MACAHY VRRIEENKGM
Dolphin I   .....EKTEI PAVF..NFLN .......... SDYTNYLLFC ME......VS ----LTCAYL ARTLQVDDGV
Donkey      KKIFAEKTES PAEFKINYLD EDTVF--ALD SDYKNYLFLC MKNAATPGQS ----LVCNYL ARTQMVDEEI
Horse I     KKIFAEKTQS PAQFKINALD EDTVF--YLD TDYKNYLFLC MKNAATPGQS ----LVC-YL ARTQMVDEEI
Horse II    RNIVAQKTES PAEFKTEDSA VFTVN--YQP GERKNYLFLC MKDVGPCLPS AEHGMVCQYL ARTQKVDEEV
Pig I       EVLLAKKTDI PAVFKINALD ENQLF--LLD TDYDSHLFLC MENASQ-EHS ----LVQQCL ARTLEVDDQI

             * *        *                  *
Bovine A    LEKFDKALKA -LPMHIRLSF NPTQLEEQCHI
Kangaroo    NE-FKKILRT -LAMPYTVI- -EVRTRDMCHV
Dolphin I   MEKFNKAIKP ALPMHIR-FS PTQLEE.....
Donkey      MEKFRRALQP -LPGRVQILP DLTRMAERCRI
Horse I     MEKFRRALQP -LPGRVQIVP DLTRMAERCRI
Horse II    MEKFSRALQP -LPGRVQIVQ DPSGGQERCGF
Pig I       REKFEDALKT -LSVPMRIL- -PAQLEEQCRV
```

Key

.. residues not yet determined
- gap inserted manually to help alignment
* indicates a residue conserved in the above sequences

by a variety of methods. Osmotic pressure [87], sedimentation-diffusion [88] and light-scattering [89] experiments all gave values around 36kDa. However BLG, when on ammonium sulphate films [90], reducing SDS-PAGE gels, or studied by amino acid analysis [70] revealed a molecular weight of 17-18.5kDa. This suggested that bovine BLG was a dimeric protein capable of existing as a monomer under certain conditions. The effect of pH on its state of aggregation has been studied.

Over the pH range 1.6 to 3.5 bovine BLG has been shown, by sedimentation equilibrium experiments, to undergo a rapid, reversible interconversion between monomer and dimer [92]. Light-scattering experiments have shown that the magnitude of the attractive energy is consistent with the presence of hydrophobic bonds, whilst the decrease in the dissociation constant ($2.5x10^{-4}$mol/l at pH 1.6 to $4.3x10^{-6}$mol/l at pH 3.5) implies that electrostatic interactions are important [92].

Between pH 3.7 and 6.5 bovine BLGs are predominantly dimeric; although BLG-A, and mixtures of A and B, can reversibly octamerize to produce aggregates of 144kDa [94,98]. Neither tetrameric nor hexameric intermediates are present in significant amounts. The interactions within the octamer of BLG-A are maximal around pH 4.5 and near 0^{o}C. The packing of subunits to form this compact octamer with 422 symmetry [99] is accompanied by an increase in hydration - proton NMR has indicated that 5-6 extra water molecules/dimer BLG are present in the octamer [97]. Aromatic residues are not believed to be involved as their UV spectra were unaltered by this association. However the free thiol, Cys-121 [150], and some carboxyl groups have been implicated in this aggregation [100]. Up to 30% modification of the carboxyls by a carbodiimide had no effect on the optical rotatory dispersion(ORD) spectrum of the protein, but did reduce its ability to octamerize. Two of the carboxyl

residues likely to be involved are Asp-64 and Asp-28. The former is present in BLG-A but not BLG-B, the B variant being unable to form a pure octamer; whilst the latter is the only amino acid substituted in BLG-Dr compared to BLG-A, but enables BLG-Dr to be glycosylated at this site, which prevents octamerization, possibly due to steric hindrance. The location of these residues in the structure should enable us to determine whether they are close enough to be part of a single association site, and to suggest how the carbohydrate moiety could interfere with octamerization.

Ruminant BLGs are all believed to be dimeric over the pH range 5 to 8, although there is evidence that a monomer-dimer equilibrium exists in cow and red deer BLGs at pH 6.5 [101]. It has been proposed that the decrease in sedimentation coefficient across this pH range for bovine BLG-A and BLG-B may be linked to an increase in the dissociation of the dimer [102,103], although one report contradicts this [6]. Preliminary crystallographic studies on the dimeric form of bovine BLG have suggested that the interface between subunits involves both hydrophobic interactions between Ile-29 and Ile-147, and the stacking of the symmetry-related His-146 residues [8]. However this report failed to address the question of whether residue 28 was accessible for octamerization in BLG-A, or glycosylation in BLG-Dr. The more detailed structure presented in Chapter 2:8 should enable us to answer this question {Chapter 5:2}.

Over the pH range 8 to 9.5 slow time-dependent changes occur in BLG. The first stage of the reaction is believed to be reversible, and associated with a decrease in the sedimentation coefficient to a value representative of a monomeric structure [92]; whereas the second stage is irreversible, producing aggregates which are insoluble at the isoelectric point [105]. The addition of a thiol-blocking group inhibits aggregation,

implicating thiol oxidation and/or thiol/disulphide exchange in the formation of the heavier components [106].

1:6:2 DENATURATION.

The denaturation of bovine BLG involves the dissociation of dimer to monomer, a major change in the conformation of the polypeptide chain, and aggregation. The denaturant can be alkali, heat, organic compounds or heavy metal ions, but it is important to note that each of these may act in a different way to yield insoluble aggregates.

(a) Alkali Denaturation

Alkali denaturation affects the structure of bovine BLG in solution at pHs above 8. As the pH rises from 8 to 9 the sedimentation coefficient decreases to a value consistent with a monomeric structure [92], and then remains constant up to pH 10. Above this pH further structural changes occur: ORD studies have shown that around pH 11 the α-helix remains intact whilst the β-sheet has been converted to a random conformation [109]; whereas Fourier transform infra-red spectroscopy revealed that at pH 12 the structure was completely random [110]. These observations were made using fresh solutions, as alkaline denaturation is time-dependent, the rate of denaturation rising rapidly with pH and leading to the formation of large aggregated components [105].

(b) Thermal Denaturation

Although the thermal denaturation of BLG is unconnected with its function in vivo, studies have been carried out which provide information on the

environments of specific amino acids. Interactions between BLG and other milk components, which may be of value in the milk processing industry, are discussed very briefly.

The thermal denaturation of bovine BLG has suggested that upon increasing the temperature from 30 to 55°C the dimer is dissociated to monomer [115,117]. At higher temperatures unfolding occurs: the thiol group becomes active and is oxidized [113], and one of the two Trp residues is transferred to a more polar environment accessible to solvent [114]. It is likely that Trp-61, which is located on a flexible external loop, is the one that becomes exposed, whilst Trp-19, which is buried at the foot of the rigid β-barrel remains in its hydrophobic environment. The effect of pH on the conversion of monomers to a denatured form has indicated that the imidazole residues are involved [115]. As His-146 is close to the α-helix, this could imply some destruction of the helical structure which might explain the large change in optical rotation observed.

The effect of heating BLG in the presence of other milk components has also been investigated. Lactose stabilizes BLG against thermal denaturation [112], by forming a browning complex which is believed to be antigenic [116]; whereas κ-casein destabilizes BLG, the enhanced rate of its unfolding being entropy-driven, and indicative of hydrophobic residues becoming exposed [112]. The interaction between κ-casein and BLG is believed to involve the free thiol, the disulphide bridges and Ca^{2+} ions [117].

(c) Organic Compounds

BLG can be denatured by organic compounds, including urea, guanidinium hydrochloride, alcohol and detergent. However only those three, urea, ethanol and SDS, which

are used in later experiments will be discussed here.

Urea denaturation is believed to occur via a two-step process. At pH 3.5, where BLG tends to be monomeric, the primary step represents a first-order reversible unfolding of the polypeptide chain to a random-coiled conformation [118]. However at pH 5.2 this step follows more complex kinetics, but upon the addition of N-ethylmaleimide(NEM), which dissociates dimeric BLG, the kinetics revert to first-order [119]. This indicates that the first stage includes both dissociation and unfolding. If the change in optical rotation observed above 4M urea concentrations corresponds to the unfolding process, then it might be possible, using lower urea concentrations, to dissociate dimeric BLG whilst maintaining its secondary structure. This was investigated in Chapter 3:4:1. During the second stage of the denaturation the unfolded protein undergoes thiol/disulphide exchange reactions (and possibly thiol oxidation), to yield irreversible aggregates, including trimers of molecular weight 54kDa, which further supports the view that BLG has been dissociated.

The effect of ethanol(EtOH) on the structure of bovine BLG has been examined, as later solution experiments involved the use of this solvent. EtOH unfolds the protein, at pH 3, in a similar manner to urea [123], although unlike with urea the reaction is not entirely reversible. The unfolded protein refolds to an α-helical conformation in the presence of EtOH, which unlike urea, does not prevent the formation of hydrogen bonds. It was also observed that the optical rotation of BLG did not vary for concentrations of EtOH below 10% [123], although subsequent workers advised against using more than 5% EtOH when wishing to study the native conformation of BLG in an ethanolic solution [160].

SDS can also affect the conformation of bovine BLG, but only at high concentrations. BLG can bind one mole of

the dodecyl sulphate anion per subunit, possibly in a hydrophobic pocket on the surface of the protein, without any change in its secondary or quaternary structure. However at higher concentrations an 'all-or-none' complex can form, with 22 moles of SDS bound to the protein [124]. Continued addition of SDS causes further associations, and a change in secondary structure is observed as the BLG residues are converted from an unordered to α-helical structure.

(d) Heavy Metals

The addition of heavy metal compounds to BLG in solution induces an irreversible conformational change, which is prevented if the protein is crystallized. This suggests that the crystalline array stabilizes the native conformation [125], and enables the isomorphous heavy atom derivatives necessary for x-ray crystallography to be obtained by soaking protein crystals with heavy metal compounds.

The denaturation of BLG by heavy metals above pH 6 has been studied using mercurial compounds, silver(I) and copper salts. Ag(I) and some Hg(II) salts like sodium p-hydroxymercuribenzoate(PHMB) cause an increase in laevorotation, up to a ratio of 2 moles of salt per dimer of BLG [126], by binding to the free thiol and enhancing dissociation. Hg(II) acetate can also show similar behaviour, although at pH 6.8 the binding of only one mole of Hg(II) per dimer of BLG also shows an increase in laevorotation, consistent with a mercurial bridge forming between the thiol of each subunit [127]. For this to be the case some rearrangement of the protein must have occurred as the thiol groups are about 25A apart in the native protein [5]. The binding of one Hg(II) per dimer has been studied [128] - its reaction is first-order, specific for the free thiol, and involves a Trp, although

no direct interaction between these residues has been proved. The crystal structure reveals that Trp-19 is the one most likely to be involved as it is considerably closer to the thiol (Cys-121 -> Trp-19 is 9.1A compared to Cys-121 -> Trp-61 which is 30.7A). The denaturation of BLG by Cu(II) occurs via two consecutive first-order reactions. The first step involves the binding of one Cu(II) per subunit [126], possibly at an amide or anomalous carboxyl site, whilst the second stage involves a conformational change of the complex [129]. The subsequent loss of copper, and the low concentration used suggests that this cation might enhance the Tanford transition.

1:6:3 SECONDARY STRUCTURE.

The secondary structure of BLG in solution has been studied by a variety of techniques. Raman spectroscopy showed a strong band at 1242cm^{-1} corresponding to a combination of β-sheet and random coil conformations, and a weak band at 938cm^{-1} from which it was deduced that the protein possessed about 10% α-helix [130]. ORD studies have confirmed an α-helical content of around 10%, and suggested that the remaining 90% was divided evenly between β-sheet and random-coil [109]; whilst circular dichroism(CD) experiments gave values for each conformation of 10% α-helix, 43% anti-parallel β-sheet and 47% unordered [132]. These values agree with data obtained from infra-red spectroscopy [133], from some secondary structure prediction algorithms [134,135], and from the 2.8A resolution crystal structure, which revealed 51% anti-parallel β-strands and 7% α-helix [8].

1:6:4 CONFORMATIONAL CHANGES.

CD, and the related technique of ORD, provide an

excellent means of monitoring conformational changes. ORD and proton-binding experiments have shown that bovine BLG undergoes three pH-dependent conformational transitions which can be summarized as:

State	Q ⇌ N ⇌ R → S			
pH	4	6-6.5	7.8	>8.5

The first transition between pH 4 and 6 is reversible, and has been described for the bovine BLG variants A, B and C [104]. The observed increase in sedimentation coefficient with pH, correlates with a contraction of the protein. BLG-A undergoes a two-group ionization, whereas BLG-B follows a single proton transition, suggesting that the additional carboxyl in BLG-A, Asp-64, may be one of the groups ionized. BLG-C also follows a single proton transition, but unlike BLG-B, one extra cationic residue per subunit, possibly the additional His-59, is exposed upon increasing the pH. From ORD studies it was believed that no aromatic residues are involved in this transition [109], although one report has stated that below pH 4.5 one Trp does undergo a conformational change - this seems more reasonable as His-59, Trp-61 and Asp-64 are located in a similar area of the protein.

Between pH 6.5 and 7.8 the second reversible conformational change (N <=> R), often called the Tanford transition, is observed [6]. In bovine BLG this transition can be detected by a change in optical rotation ($[\alpha]_D$ is -25^o at pH 6, but -48^o at pH 8) [6], indicating an alteration in the environments of the optically active groups, or by a decrease in the sedimentation coefficient (3.2×10^{-13}s to 2.8×10^{-13}s) [6] {Figure 3:1}. The latter may be due to an expansion in the volume of the protein, a variation in the shape of the protein [6,104] or the increased dissociation of

dimer to monomer [95,102,103,126]. Since this transition may be of importance for the function of BLG in vivo, as it occurs over the physiological pH, it has been studied in more detail in Chapter 3.

The amino acids responsible for this conformational change are still unclear, although the effect of this transition on some residues is known. Upon increasing the pH one buried carboxyl per subunit becomes exposed and ionized, with a positive enthalpy which suggests that the buried carboxyl was originally hydrogen-bonded [137]. The pK of this anomalous carboxyl is about 7.3, which is similar to that of a His residue. Cu(II), which is known to alter the Tanford transition [126], can bind to both carboxyl and imidazole, thus making it unclear whether a His is involved. Absorption spectroscopy has shown that a Tyr is involved - it is believed that a hindered Tyr becomes partially exposed [126]. The free sulphydryl also becomes more accessible when the pH rises from 6 to 8, as demonstrated by an increase in the second-order rate constants for the binding of sodium p-chloro-mercuribenzoate(PCMB) [138,139] or tetracyanoaurate(III) [140] to bovine BLG.

The third, irreversible, conformational change is the alkali denaturation of BLG, which has been discussed previously {Chapter 1:6:2}.

1:6:5 AMINO ACID ENVIRONMENTS.

Many investigations have been carried out to determine the environments of specific amino acids within this protein.

The free sulphydryl is believed to be Cys-121, evidence coming from both crystallographic studies [8,59,141] and some enzymatic digestion experiments [142,143,144]; whilst the other four cysteine residues form two disulphide bridges, Cys-106 -> Cys-119 and Cys-

66 -> Cys-160 [145]. It has been suggested that the disulphide bridge from Cys-106 may interchange between positions Cys-119 and Cys-121 [146,147]. This proposition came from covalently binding ^{14}C-iodoacetamide to the free thiol in the presence of 8M urea, reducing the disulphide bridges, and then binding unlabelled iodoacetamide to the newly generated sulphydryls. Subsequent peptide digestion and analysis showed that the label was evenly distributed between Cys-119 and Cys-121 [147]. However it should be noted that there was no appreciable reaction between the free thiol and iodoacetamide in the absence of 8M urea [146] which is known to disrupt disulphide bridges [147]. Hence the 'interchange' observed could be explained by urea breaking the Cys-106 to Cys-119 disulphide bridge and unfolding the protein so that both Cys-119 and Cys-121 become equally exposed and Cys-106 is buried. Preferential crystallization of one of the two equally abundant forms of the native protein also seems unlikely as much more than 50% of the BLG in a crystallization tube crystallizes, but only the presence of Cys-106 -> Cys-119 is detected. Site-directed mutagenesis has recently been used to replace each of the residues Cys-119 and Cys-121 in turn by Ser [93]. The mutant proteins have not yet been analysed but should resolve the anomaly over the position of the free thiol.

The environment of the free thiol has been investigated using a variety of compounds. The rate of binding of PCMB increases above pH 6 (by a factor greater than that due to the ionization of the -SH), indicating that the thiol becomes more exposed at higher pH [138] - but whether this is due to the conformational change or the dissociation of the subunits is unclear. It is known that when NEM binds to the free sulphydryl of BLG the protein is completely dissociated [127], implying that the thiol is located near the site of association.

However some modifications of the free sulphydryl can yield derivatives with an unaltered ability to dissociate from dimer to monomer [149,150]. From a comparison of the van der Waals dimensions of the different substituents it has been suggested that the free thiol and association site must be at least 6A apart. The distances between the free thiol and residues involved in the subunit association site are presented in Chapter 5:2:2.

There are four tyrosine residues per subunit of bovine BLG-A and their accessibility has been investigated by reactions with N-acetylimidazole and cyanuric fluoride [150]. Two of the Tyr are readily accessible at pH 9.3, and one can be acetylated with no change in the proteins CD spectrum; whereas the third is partially hindered and only becomes reactive at pH 10.8 after some conformational change of the protein, and the fourth is unavailable. None are believed to be involved in subunit association, although one or more are affected by the Tanford transition [126]. Which Tyr residues exhibit which properties should become clear from their environments in the crystal structure, and a comparison of the lattice X and Y structures.

Bovine BLG contains two tryptophans per monomer, Trp-19 and Trp-61. Trp fluorescence studies have suggested that both residues are in a hydrophobic environment at 20^{o}C [114], whilst solvent perturbation spectroscopy revealed that about 50% of these residues are buried. This is consistent with the crystallographic structure in which Trp-19 is buried at the foot of the hydrophobic cavity and Trp-61 is in an ill-defined flexible loop. The modification of BLG by HNBB in the absence of 6M urea confirmed that one Trp per subunit was accessible [155]; whereas in the presence of 6M urea both Trp residues were affected, and the UV spectrum of the BLG+retinol complex shifted to that of free retinol [153]. The exposure of Trp residues decreases slightly

above pH 6 [150], implying that they may be involved in the Tanford transition. It seems probable that Trp-61 is not involved in the Tanford transition, as electron transfer between Trp and Tyr amino acids is predominantly initiated by the electron-deficient Trp-61, and its rate is independent of pH over the range 6.1 to 7.9 [152]. The Trp environments are also slightly affected by the state of association, indicating that either Trp-61 is near the site of association or Trp-19 is involved. The latter seems unlikely as this residue is buried within the β-barrel of the protein.

Phosphorylation of BLG by $POCl_3$ in CCl_4 produces no phosphoserines, the major species being N-phospholysine and/or N-phosphohistidine [154]. The extent of phosphorylation increases with pH, but never exceeds the stoichiometry of 14 moles phosphate/mole of BLG. This is less than the total number of Lys and His residues (15+2) indicating that some are not fully exposed. Most of the Lys residues in bovine BLG are accessible for reductive alkylation although those in the α-helix (Lys 135, 138, and 141) are only partially modified, and Lys-47 remains unaltered [155]. Cleavage of the penultimate His and C-terminal Ile of bovine BLGs has been carried out using carboxypeptidase-A [156]. Both UV spectroscopy and CD showed that the native and modified proteins had very similar conformations, indicating that His-161 is situated externally. The modified protein was crystallized at pH 6.5, and its 6A structure determined [5] - this confirmed the external location of His-161.

The carboxyl groups in BLG have also been studied. Treatment of bovine BLG with carbodiimide has shown that modifying 7-8 of the carboxyls prevents octamerization of BLG-A [100]. Esterification with MeOH, EtOH and BuOH affects 23, 19 and 12 carboxyls respectively out of the 25 candidates per subunit [157]. Of the two hindered carboxyls not modified by MeOH, one is likely to be the

anomalous carboxyl associated with the Tanford transition. A comparison of the crystal structures at each end of this transition should enable us to identify the anomalous carboxyl.

1:7 BINDING STUDIES.

1:7:1 SMALL MOLECULE BINDING.

A wide variety of compounds are known to bind to BLG, and some of these are listed, together with their association constants in Figure 1:12. However it is still unclear which regions of the protein are involved, and whether any of the ligands which bind are important for the biological function of this protein.

(a) Retinoids

From the structural homology observed between BLG and RBP it was suggested that BLG may serve as a retinol (vitamin A) carrier in milk [158]. As retinol is a fat soluble vitamin present in milk, one might propose that BLG could bind the retinol within its large hydrophobic cavity, thus protecting it from oxidation, and transport it - via an aqueous medium - to an appropriate receptor. Preliminary studies support this hypothesis. HPLC analyses have indicated that most of the vitamin A in fresh cows milk is associated with the protein fraction containing BLG [159]. UV absorption spectroscopy has shown that BLG and retinol form a complex [160], whilst CD has revealed that one retinol is bound per BLG subunit with a dissociation constant of 0.02μM [153] (as compared with 0.2μM for the retinol+RBP complex. The BLG complex enhances the fluorescence lifetime of retinol from 3 to 10 nanoseconds, by reducing its susceptibility to

Figure 1:12

PARAMETERS FOR LIGAND BINDING TO BOVINE BLG.

Ligand	Number bound per dimer	Association constant (M^{-1})	Reference
Retinol	2	5×10^7	[153]
Stearate	2	1.7×10^5	[176]
Palmitate	2	6.8×10^5	[176]
Laurate	2	0.5×10^5	[176]
Oleate	2	0.4×10^5	[176]
Heptane	2	0.48×10^6	[168]
Butane	2	1.7×10^3 (5.8×10^2)	[169]
Pentane	2	7.1×10^3 (6.2×10^3)	[169]
Iodobutane	2	2.8×10^3	[169]
SDS	2	3.1×10^5	[170]
2,6-MANS	2	3.4×10^5	[162]
Methyl orange	2	0.2×10^4	[170]
n-Octylbenzene-p-sulphonate	3	6.3×10^4	[170]
p-Nitrophenol	2	1.9×10^4	[164]
p-Nitrophenylacetate	2	3.0×10^4	[164]
p-Nitrophenyl-β-glucuronide	2	1.6×10^4	[164]
p-Nitrophenyl sulphate	2	2.0×10^3	[164]
p-Nitrophenyl pyridoxal phosphate	2	3.1×10^3	[164]
2-Heptanone	2	0.15×10^3	[170]
2-Octanone	2	0.50×10^3	[170]
2-Nonanone	2	2.44×10^3	[170]
Toluene	2	4.5×10^2 (5.9×10^1)	[167]
Trifluorotoluene	2	4.2×10^2 (3.1×10^1)	[167]
Hexafluorobenzene	2	1.6×10^3	[167]

Comments

The association constants in brackets are the values obtained for the binding of a second ligand at the same site.

oxidation [161], and increases the retinol uptake by rat intestine [279]. A specific receptor for the BLG+retinol complex has also been detected in the lower third of neonate calf ileum [8].

The location and nature of the retinol binding site in BLG is uncertain. Modification of the Trp residues with HNBB caused a blue shift in the absorption spectrum of the BLG+retinol complex to that of free retinol [153], suggesting that bonding interactions have been destroyed, and implicating Trp in the binding site. The binding is independent of pH in the range 2 to 7.5, suggesting that a non-ionizable group is involved. As HNBB treatment can also affect free sulphydryls, Cys-121 may be involved in binding. Crystallographic studies of the trigonal lattice Z crystals, grown at pH 7.5 from a solution containing bovine BLG and excess retinol, have shown that retinol apparently binds in an external hydrophobic channel, adjacent to the α-helix and Cys-121, and not in the hydrophobic cavity near Trp-19 [59]. The relative importance of these two possible binding sites for retinol will be discussed, in the light of additional knowledge gained from our later binding studies, the behaviour of other proteins homologous to BLG, and crystallographic studies.

(b) Aromatic Compounds

Fused and single ring aromatic compounds are known to bind to BLG. The binding of 2,6-MANS has been studied by calorimetry, equilibrium dialysis and fluorescence spectrometry, and reveals that bovine BLG-B possesses two anionic binding sites, one of which binds more strongly than the other [162]. Both of these sites exist at each end of the Tanford transition, are believed to be hydrophobic, and may correspond to the two sites suggested for retinol binding.

Single ring aromatics also bind to BLG. The hydrolysis of p-nitrophenyl phosphate(PNPP) by bovine spleen phosphoprotein phosphatase was partially inhibited by BLG suggesting that BLG could bind PNPP [163]. The binding of several p-nitrophenyl compounds to bovine BLG-A was confirmed by Trp fluorescence studies [164], implicating Trp in the binding. However the dissociation constant for the 1:1 complex of bovine BLG and PNPP was independent of pH over the range 4 to 7.5 [164], indicating that this binding was not affected by the Tanford transition.

The binding of toluene, trifluorotoluene and hexafluorobenzene to bovine BLG has been investigated, and reveals two distinct forms of binding : strong association at a single hydrophobic site per subunit with an accessible volume of $330A^3$ (which is believed to be the alkane binding site), and weaker interactions at one or more other sites. Neither type of association was affected by dimerization of the protein [167].

(c) Aliphatic Compounds

Many alkanes can bind to BLG, the amount bound decreasing as the chain length increases [168]. Butane, pentane and iodobutane all bind to a single hydrophobic region per subunit of bovine BLG. The site, believed to be a hydrophobic cleft going from the surface towards the interior of the protein, is unaffected by the state of association of the protein. This cleft can bind either two butanes equally well, two pentanes unequally - possibly due to steric strain or part of the hydrocarbon tail being exposed to solvent, or one iodobutane [169]. The position of the latter, when soaked into BLG crystals and studied by x-ray crystallography, should enable us to locate the alkane binding site.

Aliphatic ketones, eg. 2-heptanone, also bind at a

single hydrophobic site in each subunit of bovine BLG, with association constants of the order of $10^3 M^{-1}$. As these are of the same magnitude as those for alkane binding, ketones may also bind in the same alkane binding cleft [170].

SDS has also been shown to bind to bovine BLG. The first stage involves the tight binding of two anions per dimer, probably at the alkane binding sites [167], to yield a complex which has been crystallized [174], but not yet studied by x-ray crystallography. Chemical modifications of BLG prior to complexation revealed that neither Trp, Cys, nor His-161 were essential, but did show that the internal His, His-146, was important for the formation of the BLG+SDS complex [173]. Higher concentrations of SDS caused a conformational change associated with the binding of 22 anions [124] - this complex exists in equilibrium with that with 2 SDS anions bound [124].

Free fatty acids(FFA) also bind to bovine BLG in a manner similar to SDS. They reversibly bind at a single strong binding site per subunit with an association constant of about $10^5 M^{-1}$, and induce a conformational change which allows further binding at up to 24 weaker secondary sites [176]. Binding at the primary site depends upon the nature of the ligand, decreasing in the order palmitic(C16:0) > stearic(C18:0) > oleic(C18:1) > lauric (C12:0). Further studies using palmitate and bovine BLG have indicated that the binding affinity increases over the pH range 6.5 to 8.5, suggesting a possible correlation with the Tanford transition. Spin-labelled stearate derivatives eg. 12-doxyl stearic acid have been studied by ESR and show some weak binding to BLG, but do not form 1:1 complexes [178]. This suggests that modifications to the alkyl chain may prevent binding at the primary site. It is interesting to note that the composition of FFAs, with more than fourteen carbons,

which binds to BLG resembles that found in milk. Whether this is coincidence or because BLG has a role as a general FFA carrier, possibly in the lipid metabolism of the mammary gland is unclear [177]. Triglycerides can also bind to BLG but no details were given [179].

Phosphatidylcholines are unable to bind to native BLG [180], but in the presence of a helix-inducing solvent BLG can refold to a structure containing about 50% α-helix, which is then able to bind to dipalmitoyl phosphatidylcholine. The vesicles formed from the sonication of the complex were studied and revealed that the ratio of lipid:BLG was 20:1, in keeping with both the SDS and FFA complexes formed. The BLG in this complex exhibited a 10% increase in Trp fluorescence, as compared to 8% when a FFA bound to BLG, although the hydrophobicity of the residues remained unaltered, indicating a reduction in quenching [180]. The ability of partially unfolded BLG to bind to lipids may be of importance for its transport across a membrane.

(d) Ions

There is evidence for the binding of both monovalent and divalent [186] cations to bovine BLG. Na^+ was able to bind to carboxyls and imidazoles in this protein above its pI value, the number of ions bound per subunit increasing from 0 to 2 as the pH rose from 6 to 9 [181]. The binding curve is of a similar shape to that observed by optical rotation over the Tanford transition. Various heavy metal cations have also been shown to bind: Ag^+ [126], Cu^{2+} [129] and mercurial salts [128], and they can all form complexes with one cation bound per free sulphydryl. The binding of mercurial salts has already been discussed {Chapter 1:6:2}.

One type of anion which is known to bind to BLG is the iodide ion. I_3^- reacts specifically with the free

thiols of bovine BLG producing a sulphenyl iodide complex. This is less stable than the native protein and is able to react with different mercaptans to yield derivatives with varying abilities to octamerize [182].

1:7:2 MACROMOLECULE BINDING

(a) Non-milk Proteins

Two non-milk proteins, cytochrome-c(CYT-C) and renin, have been reported to interact with BLG. The initial binding of CYT-C to subunits of bovine BLG produces a 1:1 complex. If the pH is greater than 7.5 then the conformation of BLG in the complex is altered, enabling electron donation from the free sulphydryl, Cys-121 of BLG, to reduce the iron in CYT-C [183]. A non-reduced complex is known to form between CYT-C and pig BLG, the latter having no free thiol [183]. It is unclear at present whether the conformational change is linked with the Tanford transition which increases the exposure of the thiol, or which amino acids are involved in the interaction. Bovine BLGs A, B, and C all form a similar complex with CYT-C, although the rate of iron reduction depends on the variant [183].

The inhibition of human renin by bovine BLG-B has been reported [184]. It is non-competitive, and believed to require a Leu residue near a hydrophobic region of BLG.

(b) Milk Proteins

There are numerous interactions between milk proteins, excluding those formed upon heating which are outwith the scope of this thesis. Bovine BLG is known to interact with α-lactalbumin [185], and several caseins. It can react, via a BLG trimer, with κ-casein, to form a

3:1 complex involving hydrophobic interactions. This complex is subsequently stabilized by covalent bonding, and a conformational change which makes the disulphide bridges less susceptible to attack [188]. Gel filtration and affinity chromatography have recently shown that lactoferrin can also bind to BLG, forming a 2:1 complex in which ionic interactions are important [187].

(c) Immunoglobulins

Antibodies have been raised against bovine BLG-B and these cross-react with ruminant BLGs to different extents, depending upon both the species from which the BLG originates and the immunochemical method used [189]. It has been suggested that the few amino acid substitutions between ruminant BLGs are likely to lie in the antigenic regions of the protein. As monomeric BLGs show much more primary sequence variation it is not surprising that higher titres of antibovine BLG antisera are required to cross-react with pig, horse [83b], or camel [18] BLGs. This cross-reactivity suggests some degree of structural similarity between monomeric and dimeric BLGs. However it should be borne in mind that a lack of cross-reactivity with antibovine BLG antisera does not exclude the existence of monomeric BLG.

Although antibodies to bovine BLG can be used to detect BLG in the milks from various species, they are also useful for monitoring the presence of BLG in the gastro-intestinal(GI) tract, and identifying the allergenic and antigenic sites of BLG.

When cows milk was ingested by guinea-pigs the concentration of BLG in the GI tract decreased from duodenum to ileum, probably due to increased degradation by proteinases [200]. A few days after stopping the intake of cows milk the concentration of antibodies specific for BLG was found to be constant throughout the

GI tract, and tentatively identified as secIgA [201]. secIgA specific for BLG has also been detected at low levels in the saliva of infants, but its origin is unclear [199].

Some human babies, when fed cows milk (which, unlike human milk, contains BLG), show symptoms of hypersensitivity. Challenges with isolated milk proteins confirmed that BLG was the offending agent [193], whilst increasing the amount of cows milk ingested caused an increase in the amount of IgG in serum [194]. It has been suggested that the immature nature of the neonate intestinal surface enables BLG to be absorbed, and this then complexes with the IgG [197]. Why only some infants are susceptible is unclear, although the nature of the ingested BLG may be important - a monoclonal IgE antibody specific for BLG is able to recognize both native and heat-aggregated BLG, but only forms an immunocomplex with the latter causing anaphylaxis [202].

The allergenic sites of BLG are not known. The ingestion of bovine BLGs A and B by guinea-pigs led to an increase in IgG in serum, the A variant causing a greater response, whereas the injection of both BLGs A and B into sensitized guinea-pigs gave similiar hypersensitivity reactions. This indicates that the allergenic site is the same in both variants, but distinct from the antigenic site which binds IgG. The latter site probably involves the two substituted amino acids, Asp-64 and Val-118, of the A variant [203]. The extent of IgE and IgG anti-BLG antibody binding, to reduced and carboxymethylated BLG, has indicated that a disulphide bridge is not part of the sites required for allergenicity, although it is important for an antigenic region [206].

Various fragments and derivatives of bovine BLG were exposed to IgE and IgG antibodies raised against native BLG, so as to determine their antigenic activity and gain information on the antigenic sites of BLG. Tryptic,

chymotryptic and peptic digests of BLG gave no antigenic activity suggesting that the antigenic sites are discontinuous [205]. There are believed to be four antigenic sites in native BLG at both pH 6 and 8.5, indicating that neither the Tanford transition nor the state of association significantly affects the antigenic nature of the protein [208]. Modifications to the disulphide bridges, free thiol, Trp, Arg and amino groups of BLG revealed that its reaction with IgE depends upon the conformation being maintained by the disulphide bridges, but was not affected by changes to the free thiol, either Trp, two of the three Args or the amino groups [205]. Detection by IgG antibodies gave similar results, although it was now proposed that Trp and amino group moieties were also involved [208]. Additional modified BLGs were studied and also implicated His, carboxyl and possibly Tyr groups in the antigenic sites of BLG [209]. It is important to remember that chemically modified BLGs tend to have altered structures (as detected by CD [209]), and this may explain the discrepancies between results. The positions of the above amino acids in the structure of BLG should enable us to resolve the discrepancies, and locate the four probable antigenic regions in BLG {Chapter 2:8:4}.

1:8 CRYSTALLOGRAPHIC STUDIES.

Many of the methods used to isolate BLG from milk produce an impure protein, which can be purified by recrystallization. Microcrystals of BLG have been obtained during the preparation of BLG from the milks of ruminants eg. cow [1,57], goat [35], and sheep [36], and of non-ruminants eg. pig [65]. These were all too small (<0.2mm) or too disorderd to be suitable for x-ray crystallography. A variety of crystals suitable for this

type of structural study have been grown from bovine and buffalo BLGs and characterized [4,210].

Detailed x-ray crystallographic studies have been carried out on three crystal forms of bovine BLG-A which were obtained by salting-out with ammonium sulphate. Crystals of lattice X (triclinic P1 : 37.8A x 49.6A x 56.6A, 123.42°, 97.28°, 103.66°) were grown from 4M ammonium sulphate at pH 6.5; whereas crystals of lattices Y (orthorhombic $B22_12$: 55.7A x 67.2A x 81.7A, 90°, 90°, 90°) and Z (trigonal $P3_2\,2\,1$: 54.4A x 54.4A x 113.1A, 90°, 90°, 120°) were grown from 2M ammonium sulphate around pH 7.5. Of the two higher pH forms lattice Z crystals usually appear first and then convert to lattice Y, providing that the protein concentration is not too high [213]. From crystallographic data, collected on both native crystals and suitable isomorphous heavy atom derivatives, low resolution 6A structures were determined for all three lattices [5]. High resolution studies are being carried out on all three crystal forms: the 2.5A structure of lattice Z with retinol bound has just been published [59], and the 2.0A structure of lattice X is now available [141]. The preliminary, medium resolution, 2.8A structure of lattice Y is also known [213]. It was our aim to extend this structure to higher resolution, and then compare these three structures.

1:9 THREE-DIMENSIONAL STRUCTURES.

Low resolution 6A structures of lattices X, Y and Z have been obtained by x-ray crystallography using isomorphous derivatives [5]. All three structures show features suggestive of an α-helix and some β-sheet, but no distinct differences.

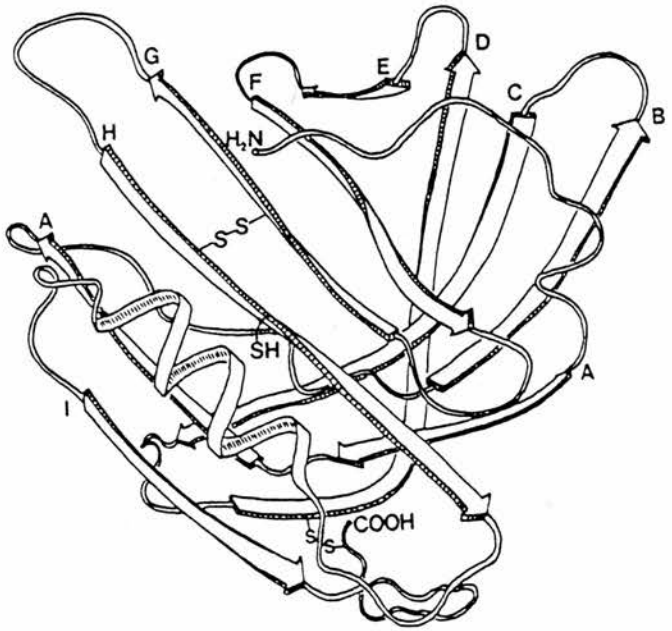
The first detailed structure of bovine BLG to be obtained was that of the orthorhombic lattice Y at medium

resolution [8,158,213]. This 2.8A structure consists of a three-turn α-helix near the C-terminus and nine β-strands, eight of which form an anti-parallel β-sheet that wraps around to create a flattened β-barrel, closed at one end {Figure 1:13}. The positions of two possible binding residues, Trp-19 and Cys-121, and the interface between subunits have been located. Trp-19 is situated at the foot of the β-barrel, on the inner surface - modelling retinol into the cavity revealed that the β-ionone ring is 10A away from the Trp, whilst its hydroxyl is near Lys-70 [8]. The free thiol, Cys-121, is located near the surface, in a hydrophobic channel bordered on one side by the α-helix. The interface between subunits is believed to involve hydrophobic interactions - this structure suggests that Ile-29, Ile-147 and the stacking of the imidazole rings of the symmetry-equivalents of His-146 could be important [8].

A comparison made between BLG and serum RBP showed only moderate amino acid identity (25%) [214,80], but a very high degree of structural homology [8]. The eight-stranded β-barrel, C-terminal α-helix and disulphide bridges were conserved in both proteins, only the lengths and positions of the loops differing significantly. The crystallographic structure of the isolated RBP+retinol complex revealed that the chromophore was located high up within the hydrophobic cavity, prevented from reaching the base by a ring of five Phe residues [215]; and this may be its position <u>in vivo</u>. Removal of the chromophore produced apo-RBP, which has not yet been crystallized, although its structure has been modelled by molecular dynamics(MD) [218]. This revealed that upon removal of the chromophore the β-barrel remained intact, and the α-helix moved towards the position it occupies in BLG. This may indicate that whatever ligand binds to BLG <u>in vivo</u> is less firmly bound than retinol is to RBP, and is thus removed by the proteins isolation procedure. <u>In vitro</u>

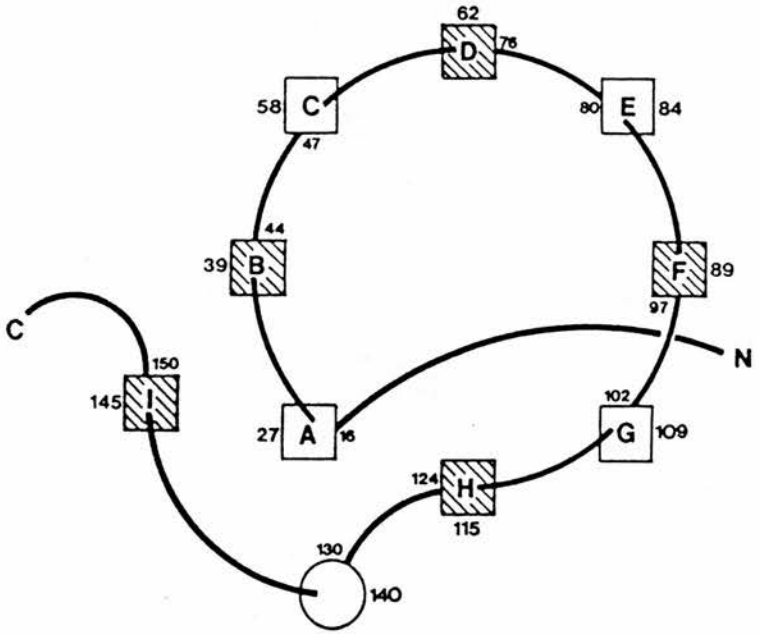
Figure 1:13

CARTOON DIAGRAMS OF BOVINE BLG SHOWING ITS SECONDARY STRUCTURE ELEMENTS.



Drawing of the BLG fold.

The arrows labelled A-I represent the strands of β-sheet.



The topology of BLG.

The N-termini of the strands are hatched to emphasise their antiparallel nature.

studies have shown that RBP can bind a variety of retinoids [216], suggesting that the binding site does not exhibit a very tight fit. This seems contrary to that expected if binding was within the hydrophobic cavity. An external binding site has been proposed from the binding of β-ionylidene ethylbromoacetate(IEBA) to RBP in vitro, and is located near the disulphide bridge Cys-4 -> Cys-160, and involves the residues 1-5 and 156-163, the beginning of the latter region being at the C-terminus of the α-helix [217]. This position appears to be similar to the hydrophobic channel of BLG into which retinol binds prior to crystallization of the complex [59]. The nature of this binding site may be partially deduced from the type of compounds which can bind, the strength of retinol binding to this site in both RBP and BLG, and the differences between these proteins upon binding a Schiff base [216].

The β-barrel structure, consisting of two orthogonal anti-parallel β-sheets, is found in two families of proteins: those with 8-stranded, and those with 10-stranded β-barrels. The BLG family of proteins have 8-stranded β-barrels and an α-helix near their C-terminus, and include RBP [215] and two biliproteins, insecticyanin and bilin binding protein. Both the INCYN structure at 2.6A resolution [219,220,222] and the BBP structure at 2.0A [221,223] are very similar to BLG, RBP and one another. They both bind biliverdin IXγ (in a cyclic helical conformation) at the same position within, but near the top of, the hydrophobic cavity, although the orientation of the chromophore differs considerably between these two biliproteins. Despite the structural homology between INCYN, BBP and BLG, it is unlikely that BLG binds biliverdin IXγ in vivo, as although biliverdin is found in buffaloes milk, it is only present at a low concentration (0.6μg/ml), is associated with the casein fraction, and exists in a different isomeric form [224].

It is unclear at present how the family of proteins containing 10-stranded β-barrels is related to the BLG family. Photoactive yellow protein has a similar structure to RBP - its two extra strands appear to be extensions of strands in RBP - and it binds a chromophore, probably a retinoid, between the β-sheets [225a]. Rat intestinal fatty acid binding protein(IFABP) and P2 myelin protein(P2), which are both members of a family of proteins that includes the cellular retinoid binding proteins, adopt the 10-stranded β-barrel structure. The IFABP structure has a large gap between two of its adjacent strands, part of which is obstructed by solvent molecules, but may still be the way in which palmitate enters the β-barrel [225b,229]. P2 also binds a fatty acid in a location similar to that of retinol in RBP, although its orientation is reversed - the hydrophilic carboxyl pointing into the barrel towards two Arg residues [225c]. The Cα traces of these β-barrel proteins, and the location of their ligands, is illustrated in Figure 1:14.

1:10 HOMOLOGOUS PROTEINS.

The primary sequences of the four structurally homologous, 8-stranded β-barrel proteins: bovine BLG-B [71], serum RBP [226a], INCYN [226b] and BBP[223] were aligned [227], and common features noted:

(a) a sequence Asn.h.a.x.x.b.h.x.Gly.x.Trp.y.x.h.Ala around residue 20 (BLG numbering)

(b) a sequence h.h.x.Thr.Asp.Tyr.x.x.Tyr.h.h near residue 100

(c) a conserved basic residue at position 124

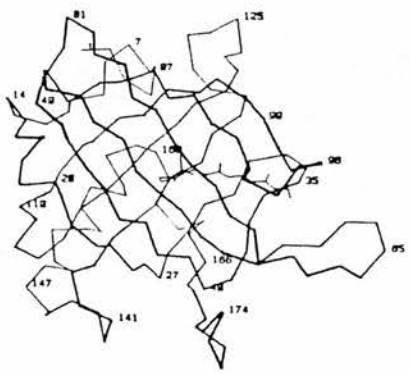
(d) Cys residues, which could form a disulphide bridge, near positions 66 and 160,

where a is often acidic, b is usually basic, h is

Figure 1:14

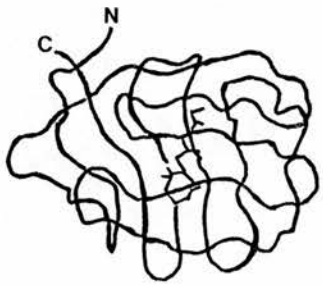
Cα TRACES OF β-BARREL PROTEINS SHOWING THE LOCATION OF THEIR RESPECTIVE LIGANDS.

RETINOL BINDING PROTEIN



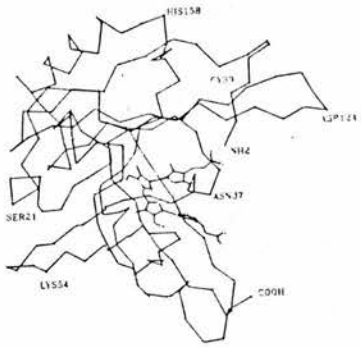
(215)

PHOTOACTIVE YELLOW PROTEIN



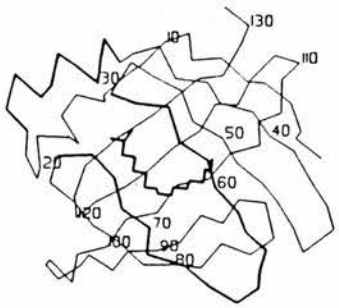
(225a)

INSECTICYANIN



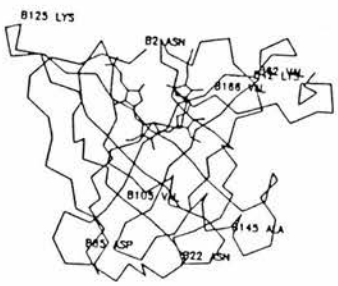
(222)

INTESTINAL FATTY ACID BINDING PROTEIN



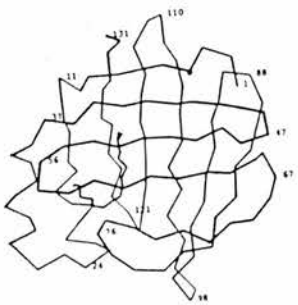
(225b)

BILIN BINDING PROTEIN



(223)

P2 MYELIN PROTEIN



(225c)

hydrophobic, y is aromatic, and x represents any kind of residue. Although these common amino acids represent only a small proportion of the sequence, they are all known to be grouped spatially at the base of the hydrophobic cavity [228]. Hence this region may be important for binding to a cell surface receptor.

These common sequences have now been identified in many proteins, and have led to the proposal that BLG belongs to a superfamily of proteins [9], recently called the 'lipocalins' [239]. All the members have molecular weights around 18kDa, are likely to adopt 8-stranded anti-parallel β-barrel structures made up of two orthogonal β-sheets, and probably bind and transport hydrophobic/labile/insoluble ligands within their hydrophobic cavities. Members of this family include other retinoid binding proteins in addition to RBP: purpurin(PURP), a neural retina cell adhesion protein which binds retinol and may transport it across the interphotoreceptor cell matrix [232]; crustacyanin (CRCYN), a carotenoid binding protein from lobster which binds the blue pigment astaxanthin - release of this chromophore upon boiling gives the lobster its characteristic red colour [231]; and α_1-microglobulin (A1UG)[233] and protein HC [234], both of which bind a brown-yellow chromophore that has in the latter case been tentatively identified as a retinoid prosthetic group [80]. Protein HC has also been shown to have 23% amino acid homology to complement 8γ [235], which contains the consensus regions, and is probably another member of this superfamily.

Cellular retinoid binding proteins are currently not believed to be part of this superfamily, as their sequences do not contain the above common regions. However it does seem likely that they will adopt the 10-stranded β-barrel structure found in P2 and IFABP [225c, 229].

Other vitamin binding proteins have had their sequences aligned with BLG. Vitamin D binding protein contains none of the consensus regions [236], whereas streptavidin (vitamin H (biotin) binding protein) possesses the first common region [237] but only shows 4% overall homology to BLG. However, surprisingly, its structure has an 8-stranded β-barrel composed of two orthogonal sheets, albeit only poorly superimposable upon RBP [238]. This, together with the knowledge that FABPs do not contain the common features of the BLG superfamily [225c] but IFABP still adopts a β-barrel structure [229], indicates that the structural motif of 'two orthogonal β-sheets forming a β-barrel' is common in transport proteins, and does not depend on the consensus features of the BLG superfamily. This agrees well with the idea that the conserved residues may be important for binding the protein to a cell surface receptor rather than to a ligand, as they are located on the surface of BLG in the same spatial region [238].

Several proteins are now known to contain the common features of the BLG superfamily, but as yet there is limited evidence as to whether their structures possess the β-barrel motif. Apolipoprotein D(APOD) contains the conserved regions of the BLG superfamily, and from its primary sequence has been predicted to have less than 5% α-helix [250,251]. Its tertiary structure has been modelled using the coordinates of INCYN [244]. APOD forms part of the 1:1:2 lecithin cholesterol acyltransferase : apolipoprotein A1 : APOD complex found in plasma which is believed to transport the cholesteryl esters synthesized by the enzyme [249]. However it now seems unlikely that APOD can bind cholesteryl esters, as the ligand is too long to fit within the β-barrel and would have to be involved in unfavourable interactions with the protein. APOD is known to form a near 1:1 complex with bilirubin IXα <u>in vitro</u>, and is now believed to have a role as a

mammalian bilin binding protein [244].

Molecular cloning techniques were used to isolate a protein from the olfactory tissues of frog which contained the conserved amino acid regions present in the BLG superfamily [253], whilst the binding of a pyrazine derivative to bovine olfactory epithelium enabled two groups to isolate bovine odorant binding protein(OBP) [254, 255]. These proteins appear similar to BLG, being dimeric, unglycosylated, having a molecular weight of 19kDa, and a pI around 4.8; although the first 68 amino acids of bovine OBP so far sequenced only show partial homology to the initial consensus region in the superfamily [258]. Partial homology is also observed in rat OBP where both main common features are modified [257]. However as these OBPs are believed to concentrate and transport odorants across the hydrophilic nasal mucosa, they may still be part of the BLG superfamily of transport proteins.

The partial sequence of the bovine OBP (cow nasal pyrazine binding protein) shows a high homology (33%) to mouse major urinary proteins(MUP) and the rat urine proteins, α_{2u}-globulins(A2UG) [258], and hence these urine proteins could also be members of the superfamily.

Sequence homology (about 20%) was observed between these urine proteins and aphrodisin, a soluble 17kDa, 151 amino acid protein. This protein contains two disulphide bridges but only the first consensus region, making its inclusion in the superfamily still rather tentative [264]. It has been extracted from hamster vaginal discharge and shown to elicit copulatory behaviour in male hamsters. However attempts to extract any pheromone from this protein have so far failed.

mRNA coding for the 20kDa MUPs has been found in several mammalian-specific secretory tissues including the liver, lachrymal gland and mammary gland [248]. The level in female liver is only 20% of that found in male

liver, and is stimulated by growth hormone and thyroxine, rather than testosterone; whereas the level in mammary gland is about 1/30 of that found in male livers, is present throughout pregnancy but decreases just before parturition.

Different forms of A2UG have been found in male and female rat urine [259]. The male-specific protein is non-glycosylated, androgen-regulated, possesses the conserved regions [260], and is believed to bind pheromones, and 2,4,4-trimethylpentan-2-ol [261]. The latter complex can be isolated from kidney and urine suggesting that A2UG is also a transport protein. Two other androgen-dependent secretory proteins(ESP) have been isolated from rat epididymis with molecular weights around 18.5kDa [262]. They differ only near their N-terminus, both containing the conserved regions and showing high homology to other members of the BLG superfamily [263].

As both male-specific proteins, ESP and A2UG, and female-specific BLG belong to this superfamily of proteins, other sex-specific proteins were examined. One of the human placental proteins, PP14, has been fully sequenced [45] and shown to contain both the homologous regions present in the BLG superfamily. This glycoprotein has been detected in seminal plasma, uterine luminal fluid during the luteal phase of the menstrual cycle and in amniotic fluid [266]. In women its synthesis and secretion may be related to the progesterone-dependent differentiation of the glandular epithelium.

α_1-acid glycoprotein(A1AG) has been isolated from serum with a lipid bound, and can, after delipidation, bind progesterone [267] or chlorpromazine [268] at a single site. A comparison of its primary sequence from rat, mouse and human [270] failed to confirm whether this protein is a member of the BLG superfamily. All three A1AGs contain two appropriate Cys residues (although it

is unknown whether they form a disulphide bridge), but only the human protein has the first main consensus region, and none have the second. However the exclusion of A1AG from this superfamily seems unreasonable as its gene structure shows remarkable similarity to those of RBP, APOD, MUP and BLG [275], and its predicted secondary structure agrees well with the known structures of BLG and RBP [239].

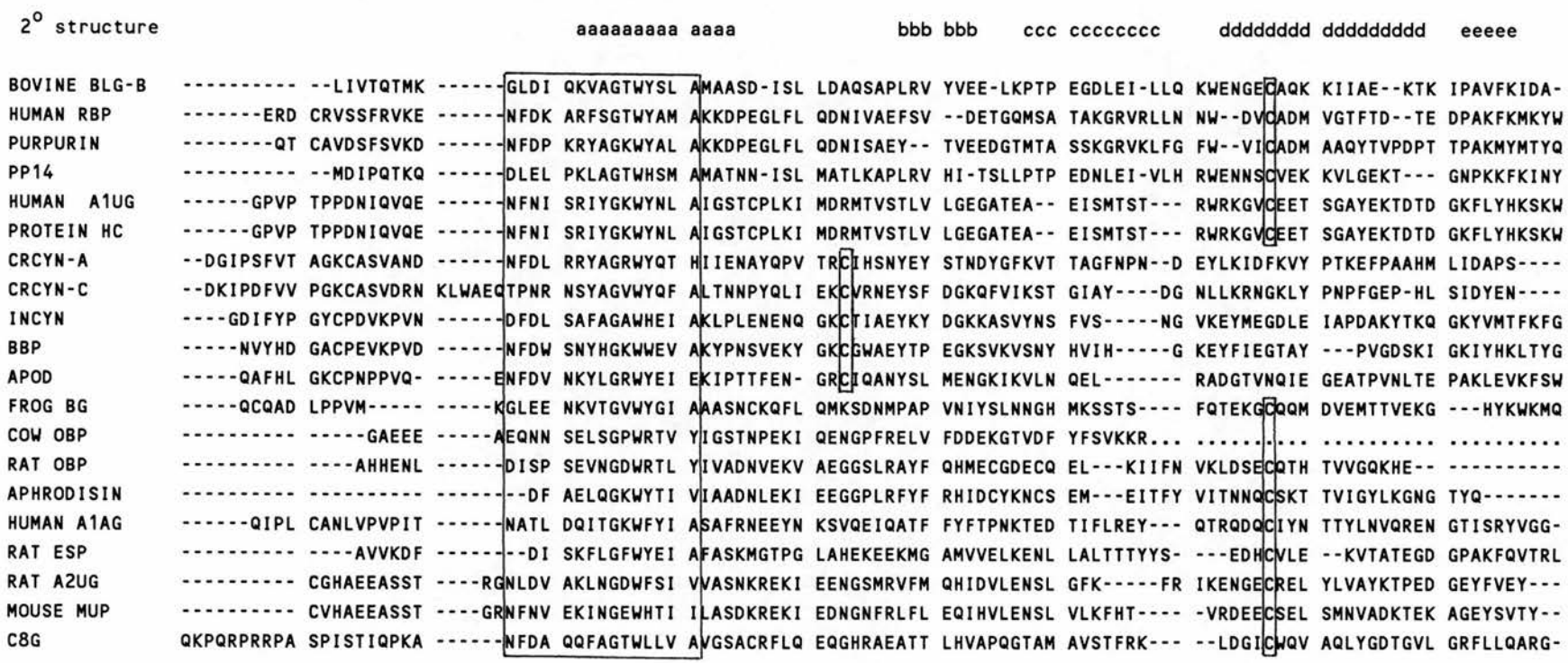
Comparative studies on the primary sequences of these proteins have been summarized in Figure 1:15. It appears that BLG belongs to a superfamily of proteins [9], which have only a few common amino acid regions, but these are located in the same spatial region, on the surface of the protein [228]. The relatedness of these proteins is further supported by the gene homology between five members of the superfamily [275]. The properties of these proteins are summarized in Figure 1:16. They all have molecular weights around 18kDa, are secretory proteins which have the ability to bind small conjugated hydrophobic/labile ligands probably within their hydrophobic β-barrel cavity (as identified in BLG, RBP, BBP, and INCYN [8,215,223,222]), and to transport them to specific receptors which have been detected for BLG and RBP [8]. It has been suggested that the invariant Trp-19 which points into the ligand binding cavity may be involved in inducing a conformational change that weakens the interaction between the unloaded protein and the receptor [272].

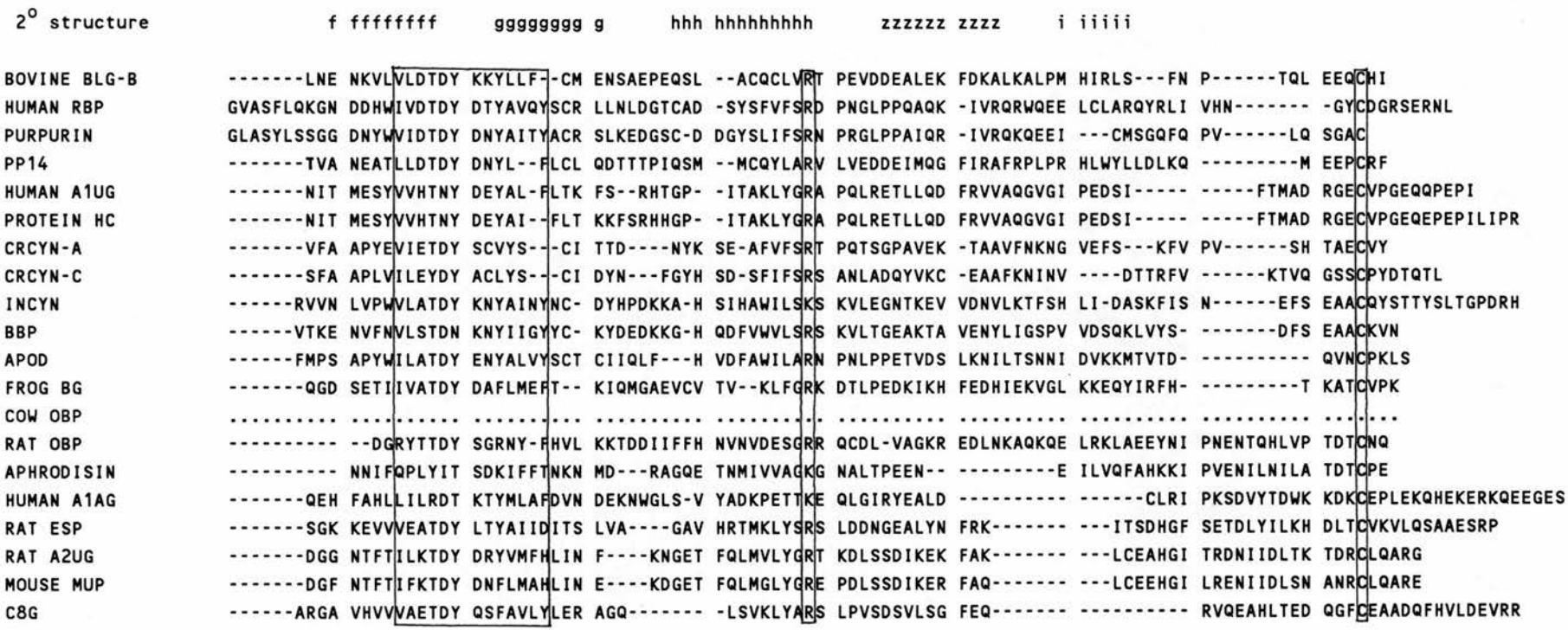
1:11 GENETIC ASPECTS.

The abundance, stability down to pH 2, and large hydrophobic cavity of BLG indicate that this protein may be of significant use for the binding, and subsequent transportation of an insoluble drug, unmetabolized, to an

Figure 1:15

PRIMARY SEQUENCES OF MEMBERS OF THE BLG SUPERFAMILY.





Comments.

The abbreviations used for the protein names, and the one-lettered amino acid code, are given in the Table of Abbreviations.

a,b,c,d,e,f,g,h,z,i represent the secondary structure elements strand A, strand B, strand C, strand D, strand E, strand F, strand G, strand H, α-helix, and strand I, as taken from [8] and illustrated in Figure 1:13.

The enclosed boxes indicate the conserved regions in this superfamily.

.. = residues not yet sequenced
- = gap put in to aid manual alignment

Figure 1:16

SUMMARY OF SOME CHARACTERISTICS OF THE BLG SUPERFAMILY PROTEINS.

Protein	Species	Number of residues	Fluid in which located	Ligand	Protein it binds to	Function	Reference
BLG	Cow	162	Milk	Palmitate, retinol	Intestinal receptor	?Fatty acid transport	[8,176]
RBP	Human	182	Serum	Retinol	Transthyretin	Transports retinol	[216,217]
PURP	Chicken	175	Interphotoreceptor cell matrix(IPCM)	Retinol	unknown	?Transports retinol across IPCM, promotes embryonic neural cell adhesion	[232]
A1UG	Human	167 }	Serum,urine	Retinoid, porphyrins	Immunoglobulin A	Mediation of neutrophil chemotaxis	[80,234]
PROTEIN HC		181 }					
CRCYN-A	Lobster	168}	Carapace	Carotenoid, astaxanthin	unknown	Colouration	[231]
CRCYN-C		175}					
INCYN	Moth	189 }	Hemolymph	Biliverdins, heme	unknown	Camoflage colouring	[226b]
BBP	Butterfly	173 }					
APOD	Human	169	Serum	Bilirubin, cholesterol	Lecithin:cholesteryl transferase, apolipoprotein A1	Localized acute phase protein	[244,249,251]
A1AG	Human	181	Serum,urine	Progesterone, lipid steroids,	unknown	Acute phase protein	[268]
FROG BG	Frog	159}	Nasal mucosa	Pyrazines, other odorants	unknown	?Presentation of odorants to receptors on neurons	[253]
COW OBP	Cow	unknown}					
RAT OBP	Rat	172}					
A2UG	Rat	162 }	Urine,serum	2,4,4-TMP, ?pheromones	unknown	?Binding of pheromenes	[260]
MUP	Mouse	181 }					
APHRODISIN	Hamster	151	Vaginal discharge	?pheromones	unknown	Elicits copulatory behaviour in male hamsters	[264]
ESP	Rat	166	Epididymal fluid	unknown	Sperm membrane protein	unknown	[262]
PP14	Human	162	Amniotic fluid, seminal plasma	unknown	unknown	Present in mid-luteal phase to early pregnancy	[52,56]
C8G	Human	182	Serum	unknown	Complements proteins	Inessential part of lysing complement complex	[235]

intestinal receptor. Modifications to the gene coding for BLG, by site-directed mutagenesis, would enable mutants to be produced that could bind different compounds.

An ovine mammary cDNA library was prepared - cDNA for BLG was inserted into the plasmid pBR322 and then cloned into E.coli. Northern blot analysis showed that the 850bp mRNA for ovine BLG was homologous to, and similar in size to that identified from porcine [273]. Two cDNA recombinant plasmids were then chemically sequenced, and revealed that ovine BLG mRNA contained 540bp coding for the 180 amino acid pre-BLG, and identified the location of the Tyr/His substitution between variants A and B at position 20 [274]. The BLG gene structure has been further characterized and shown to consist of 6 coding exons: exon 1 codes for the signal peptide and N-terminus, exon 2 for the strands A, B and C, exon 3 for the strands D and E, exon 4 for the F, G and H strands of the β-barrel, exon 5 for the 3-turn α-helix and strand I, and exon 6 for the C-terminus [275]. The intron/exon structure for the BLG gene is very similar to those for rat RBP, MUP, APOD and human A1AG [275]. It is interesting to note that the conserved amino acid regions are all located just after an intron/exon boundary.

The ovine BLG gene has recently been transcribed into mice. These transgenic mice have multiple copies of this gene per cell, and are capable of expressing ovine BLG in their milk in abundance (23mg/ml) [30].

Several bovine BLG cDNA clones have also been obtained. One contained 550bp of which only a few were coding - coding for the amino acids 109-122 [276], whereas another possessed bases coding for the residues 106-162 in the protein [277], and a third, for BLG-A, has been sequenced and shown to contain nearly all the gene, including 455bp which code for Leu-11 to Ile-162 of the BLG sequence [278]. More recently the complete

sequence of the bovine BLG-A cDNA has been published [271], and shows 91% homology, at the nucleotide level, to the ovine BLG cDNA.

1:12 THESIS RATIONALE.

There are many gaps in our knowledge of the properties, structure and function of BLGs. However, for bovine BLG, a lot of the physico-chemical properties are known, but its structure and biological function are still unclear.

As the pH of cows milk is 6.6, and the physiological pH is 7.4, the Tanford transition, which occurs over this pH range, is likely to be important for the function of BLG. Low 6A resolution structures of lattice X at pH 6.5 and lattices Y and Z around pH 7.5 failed to show any significant differences, thus necessitating the determination of higher resolution structures. The 2A structure of lattice X is now available [141], and it was our aim to obtain the lattice Y structure at a similar resolution (Chapter 2).

The probable importance of the Tanford transition in the function of BLG has led us to investigate the features of this conformational change in solution using two different BLGs: monomeric pig BLG and dimeric bovine BLG. These results were subsequently related to the molecular explanation of this transition obtained from a comparison of the lattice X and Y crystallographic structures (Chapter 3).

BLG is believed to be a transport protein, binding small, conjugated, hydrophobic molecules, like the other members of the superfamily to which it belongs. Thus experiments were instigated to determine whether this protein was a non-specific transport protein, capable of binding the ligands of the other members of the

superfamily, or a specific transport protein whose ligand *in vivo* has yet to be identified. The results of these binding studies and properties of the binding sites are examined {Chapter 4}.

CHAPTER 2

CRYSTALLOGRAPHIC STRUCTURE DETERMINATION OF BLG

To elucidate the molecular aspects of the Tanford transition (the conformational change which bovine BLG undergoes over the physiological pH), it is necessary to obtain high resolution crystallographic structures at both ends of, and if possible at various pH values across, this transition. A 2.0A lattice X structure at pH 6.5 [141], and a 2.5A lattice Z structure at pH 7.5 have been determined, but the latter cannot be extended to higher resolution as the crystals fail to diffract above 2.5A resolution [59]. It is our aim to obtain the structure of lattice Y at pH 7.8 in order to offer structural explanations for the Tanford transition and some of the properties of BLG.

2:1 PREPARATION OF CRYSTALS.

2:1:1 SOURCE OF BLG.

Both homozygous (BLG-A or BLG-B) and heterozygous (BLG-AB) preparations of bovine BLG were available commercially. Three of the heterozygous preparations were purchased: Sigma BLG (a three times crystallized and lyophilized powder No. L-0130), British Drug House(BDH) BLG (product 44064), and crystallized Pentex BLG (from Miles Laboratories Inc. Code No. 96-002). All three appeared to be equally free of contaminating proteins when run on a 15% SDS-PAGE gel and stained with Coomassie blue. However, during preliminary crystallization experiments, it became apparent that the Sigma BLG contained a considerable amount of salt (about 50%) which caused excessive precipitation in the crystallization

tubes. As the BDH BLG was in short supply Pentex BLG was used in all subsequent crystallization experiments.

2:1:2 CRYSTALLIZATION OF BLG.

The initial crystallization conditions tried were those described by M.Z.Papiz [213]. These produced many beautiful trigonal lattice Z crystals which failed to convert to lattice Y, even upon increasing the temperature from 4°C to 10°C. The crystallization medium was then modified by making it up to 1% w/v in n-octyl-β-glucoside, which helps break down hydrophobic interactions, but this only yielded larger lattice Z crystals.

Upon closer examination of the protocol an error was noticed. The concentration of BLG was too high by a factor of two, thus explaining the increased stability of the lattice Z crystals. Trial crystallizations were then set up using the modified procedure, but these also failed to produce lattice Y crystals. Small adjustments were made to the temperature, pH, protein and salt concentrations, and orthorhombic lattice Y crystals were obtained.

The lattice Y crystals were grown by the batch method. A 0.4M phosphate+2.8M ammonium sulphate buffered solution(PAS) was made by mixing

12ml	3.5M $(NH_4)_2SO_4$
2ml	3.0M $KH_2PO_4.2H_2O$ / KOH pH 8.5
1ml	distilled water

and adjusting the pH to 8 with pellets of KOH. The BLG solution was obtained by dissolving 131mg of protein in 1ml of 0.1M phosphate buffer at pH 7.5, and then centrifuging it at 13000rpm for 10 minutes to remove any undissolved material. The two solutions were then combined, and diluted in the following manner:

BLG solution pH 7.5	50μl
PAS solution pH 8	390-435μl
Distilled water	60-15μl

The amount of PAS was increased in 5μl volume intervals so that the concentration of ammonium sulphate varied from 2.18 to 2.44M, whilst the volume of distilled water was decreased such that the total crystallization volumes were constant. Each tube contained 13.1mg/ml of protein. The pH of the initial sample was checked, using a Corning pH meter 140, and found to be 7.8. The tubes were then sealed and left at room temperature.

After four weeks all the tubes contained microcrystals, some of the larger being recognized as lattice Z. In four tubes - those with concentrations of ammonium sulphate between 2.24 and 2.35M - the many small Z crystals gradually converted to a few larger lattice Y crystals, whilst the pH of the tubes dropped to 7.3. This suggests that the lattice Y crystals may be stabler at a lower pH compared to lattice Z. The morphologies of both lattice Y and Z crystals are illustrated {Figure 2:1}, typical dimensions for the lattice Y crystals grown were 0.5mm x 0.3mm x 0.2mm.

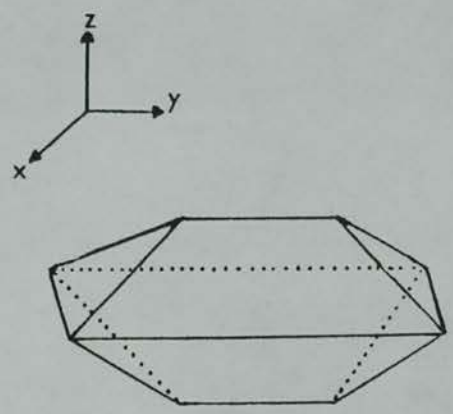
2:1:3 SOAKING EXPERIMENTS.

To investigate the binding of a ligand to BLG by x-ray crystallography, it is necessary to obtain a crystal of the complex. This can be achieved either by producing the complex and crystallizing it, or by soaking a crystal of the native protein in a solution of the ligand. There are problems associated with both methods: in the first case the binding of the ligand may alter the crystallization conditions, although this does not appear to be the case for a BLG+retinol complex which has been crystallized under the same conditions as lattice Z [59]; whereas in the second case the ligand may be prevented

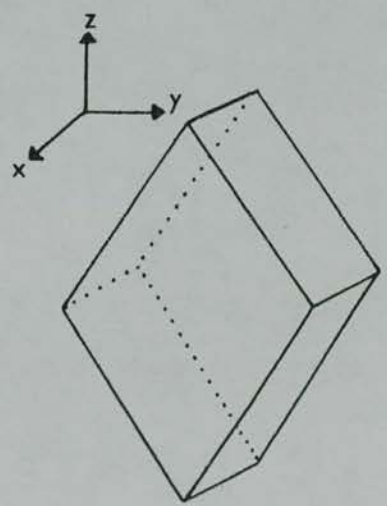
Figure 2:1

PHOTOGRAPHS OF LATTICE Y AND Z CRYSTALS OF BOVINE BLG.

Lattice Y



Lattice Z



from diffusing to the true binding site, due to its insolubility in aqueous medium or steric hindrance from the crystal packing. As the latter method is generally easier, and lattice Y crystals were available, some soaking experiments were set up.

Lattice Y crystals were removed from their crystallization medium, using a wide-ended Pasteur pipette, and transferred into mother liquor. A suitable mother liquor for BLG lattice Y crystals is 2.8M ammonium sulphate in 0.1M phosphate buffer at pH 7.7. A few crystals, in a known volume of mother liquor, were then separated off into a glass well, and a small aliquot of concentrated ligand solution was added. A variety of ligands were tried: retinoids and biliverdin (as their carrier proteins are structurally homologous to BLG), non-aqueous solvents, and both SDS and PNPP which are known to bind to BLG. The observations made are given in Figure 2:2.

Preliminary studies had shown that lattice Y crystals were unable to survive for long in highly ethanolic environments (less than 2 hours in 47.5% EtOH), and that more than 2% EtOH affected the Trp fluorescence of BLG in solution. Thus the percentage of EtOH used in these soaking experiments was kept to a minimum. Other organic solvents were also tried, and the crystals appeared to remain intact at the interface between the aqueous and organic layers. It was encouraging to see that iodobenzene appeared to cross the interface and diffuse into the crystals. This indicates that it may be feasible to dissolve a retinoid in an inert solvent and let it traverse the interface into the crystals.

The solubility of PNPP in aqueous medium should have enabled it to diffuse into the lattice Y crystals within the 72 hour duration of the soaking experiment (diffusion takes only 4 hours to get at least 90% of the ligand to equilibrium), whilst its ability to bind to BLG-A in

Figure 2:2

SOAKING CONDITIONS FOR BOVINE BLG LATTICE Y CRYSTALS IN VARIOUS LIGAND SOLUTIONS.

Ligand/Solvent	Concentration of ligand in well	Comments
Retinol/EtOH	1mM	Crystals survived for 4 hours, but were cracked by 24 hours. (EtOH content was 2%).
Retinyl palmitate/EtOH	6.3mM	2% EtOH present - the ligand remained in EtOH globules. Crystals survived in the aqueous phase for 72 hours.
Biliverdin/water	1mM	A few crystals survived for 72 hours.
Iodobutane/mother liquor	50%	Crystals stayed at interface of aqueous mother liquor and organic solvent. All crystals survived.
CCl_4/mother liquor	50%	Crystals stayed at interface of aqueous mother liquor and organic solvent. All crystals survived.
Iodobenzene/mother liquor	50%	Crystals stayed at interface of aqueous mother liquor and organic solvent. Crystals survived, and appeared to gain the orange colour of iodobenzene.
PNPP/water	10.3mM	All crystals survived for > 72 hours. Ligand solution eventually went yellow, as the PNPP was hydrolysed to p-nitrophenol.
SDS/water	10.3mM	Solution was very bitty, but crystals appeared to survive for 72 hours.

solution (K_d = 31μM [164]) suggests that a crystalline complex may have formed. X-ray diffraction data were then collected to 2.8A resolution on this complex, and the details are given in Chapter 4:7.

2:1:4 MOUNTING A CRYSTAL.

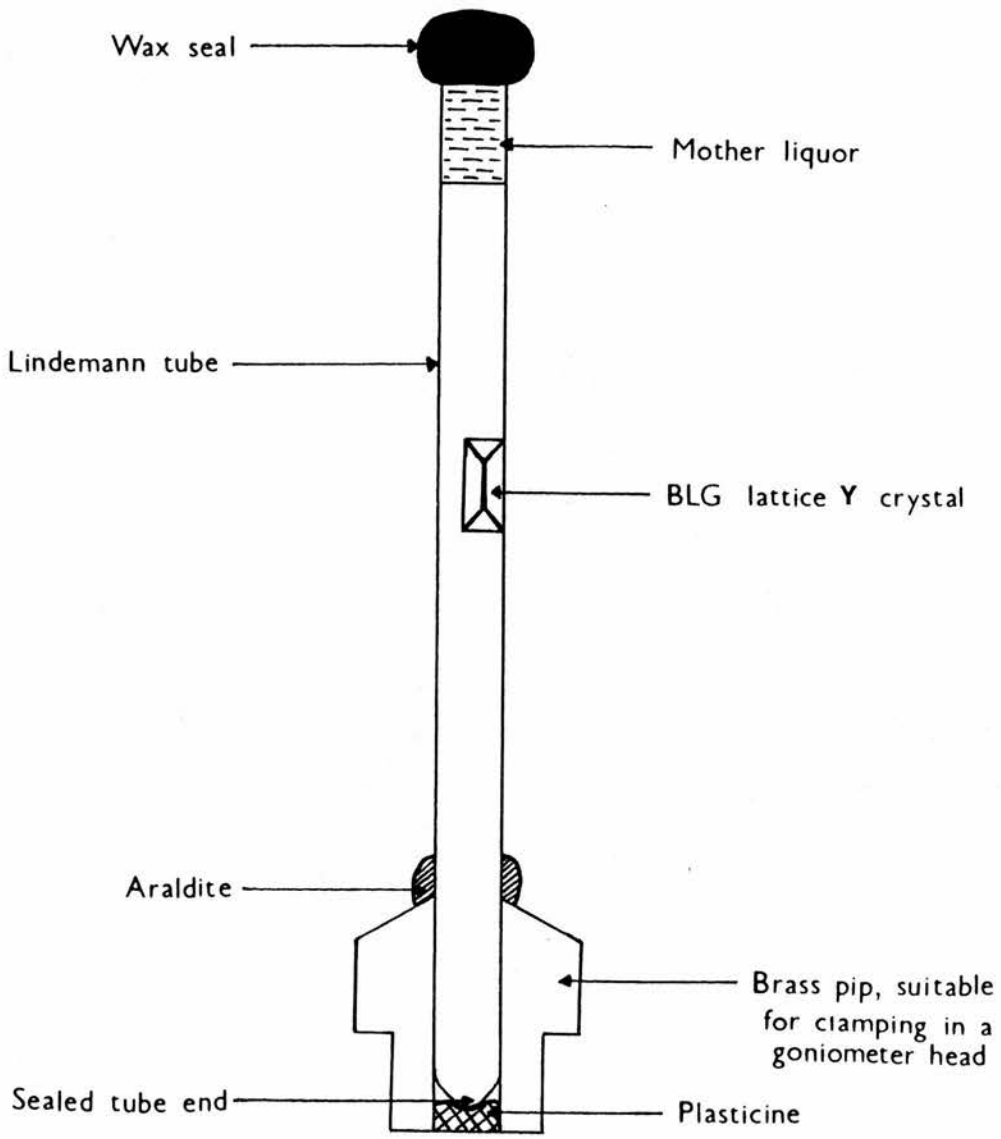
A single orthorhombic lattice Y crystal of bovine BLG was mounted in a glass Lindemann tube of diameter 1mm according to the following procedure.

The narrow end of the Lindemann tube was sealed, and firmly anchored into a brass pip, designed to fit on a Stöe goniometer head. The tube was then filled with mother liquor - from 1cm above the pip to near the top of the widened end. A single lattice Y crystal, in a drop of mother liquor, was then launched just under the surface of the liquid and allowed to sink slowly, until it settled on the meniscus. The mother liquor was removed, the capillary tube shortened, and the crystal dried. Its orientation was adjusted using thin strips of filter paper. An ideal orientation is one in which an axis of the crystal is collinear with the capillary tube axis. Immediately the crystal had been completely dried, a drop of mother liquor was replaced in the end of the capillary tube, and the tube sealed with wax. This set-up is illustrated in Figure 2:3.

The drop of mother liquor kept the enclosed environment moist and prevented the crystal drying out, whilst not aiding movement of the crystal during data collection. When cooling of the crystal was required, it was necessary to replace the drop with a short piece of filter paper dampened with mother liquor. This minimized the amount of water condensing around the crystal and dissolving it.

Figure 2:3

ILLUSTRATION OF A BLG LATTICE Y CRYSTAL MOUNTED WITHIN A LINDEMANN TUBE.



2:2 DATA COLLECTION.

When a beam of x-rays is shone on to a crystal, the electrons surrounding the atoms are made to oscillate, and these emit a secondary ray. The secondary ray is of the same wavelength, but differs in phase by 180^{o}, relative to that of the incident radiation. These scattered rays can be detected by scintillation counters, area detectors or photographic film. The theory of this scattering, and how the diffraction pattern obtained can elucidate the structure of the molecules in the crystal, is summarized in Appendix 2.

2:2:1 CALCULATIONS PRIOR TO DATA COLLECTION.

Prior to efficient data collection it was important to consider both the geometry of the camera and the symmetry of the crystal. Symbols used subsequently are defined in the List of Abbreviations.

The diffraction geometry of a rotation camera has been discussed in detail [280], and is summarized in Figure 2:4. From this geometry, and Braggs Law ($2d\sin\theta=\lambda$), it has been shown that

$$XF = R_f \,/\, \tan\{2\sin^{-1}(\lambda/2d)\}$$

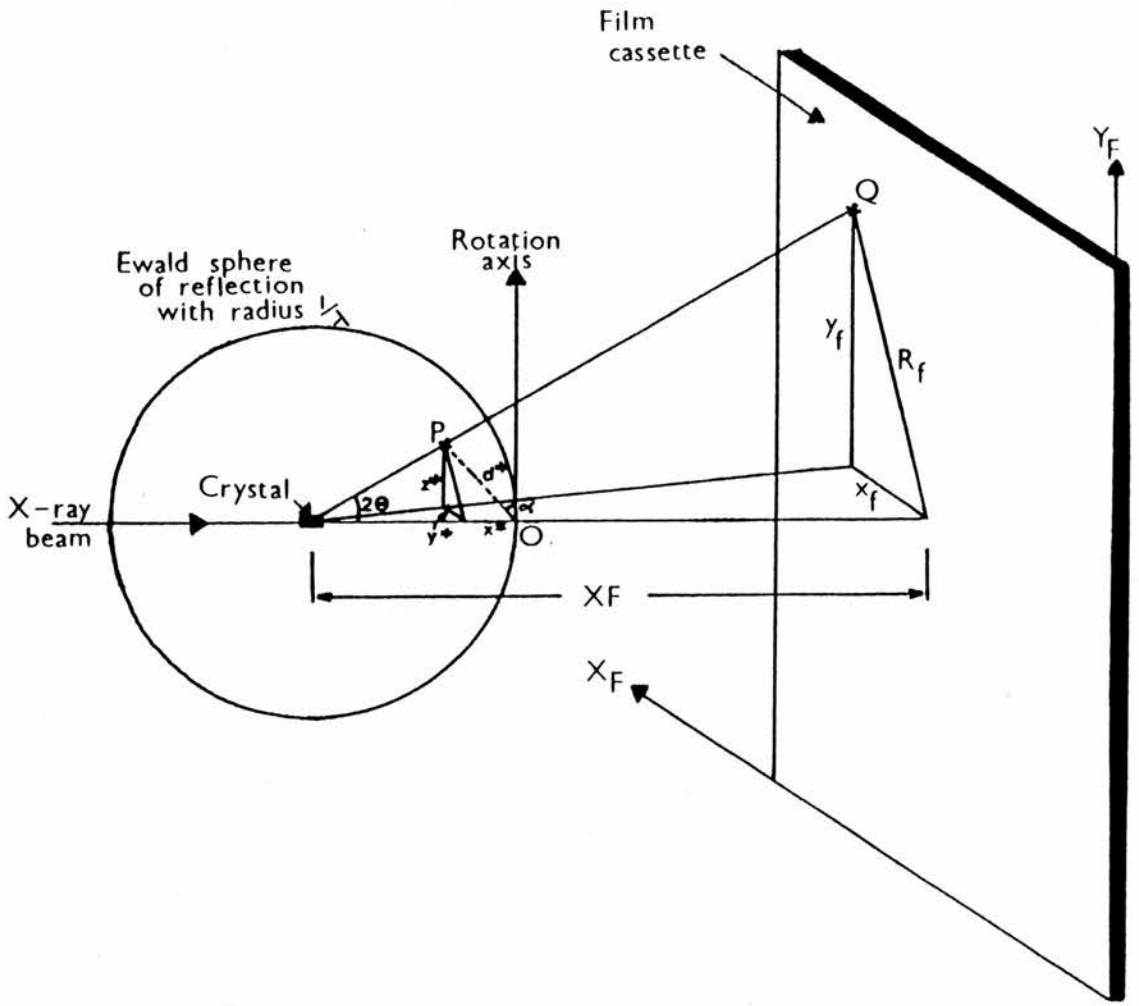
Thus to collect a 1.8A resolution dataset on lattice Y crystals of bovine BLG, using a wavelength of 0.88A and film radius of 58.5mm, a crystal-to-film distance of about 110mm was required.

The symmetry of the crystal was important for determining the rotation range necessary to collect a complete dataset. Lattice Y crystals belong to the space group $B22_12$ [5] and therefore exhibit mmm x-ray diffraction symmetry. Thus rotating the crystal through 90^{o} would cause the Ewald sphere to sweep out a portion of reciprocal space which included a unique octant of data and anomalous pairs of reflections. Some equivalent

Figure 2:4

THE DIFFRACTION GEOMETRY OF A ROTATION CAMERA.

The symbols are defined in the List of Abbreviations.



reflections would be recorded on different, non-adjacent films, and hence one could reduce the rotation range further and still collect a complete dataset {Figure 2:5}. Data were collected over more than the minimum rotation range to help improve the accuracy of the merged data. Two crystals were also mounted about a different axis, so that data from the blind region around the initial rotation axis could be measured.

The oscillation range, $\Delta\phi$, to be recorded on any one photograph must also be selected prior to data collection. An approximate value can be calculated [280] from

$$\Delta\phi = \{\lambda.d.180\}/\{1^{*}\lambda\pi\} - \Delta,$$

although it may need to be varied slightly so as to provide the best compromise between maximizing the number of full reflections recorded and minimizing the number of overlaps. Oscillation angles of 1.25^{o} to 2.5^{o} were deemed to be suitable for the collection of 1.8A resolution data on BLG lattice Y.

2:2:2 EXPERIMENTAL DATA COLLECTION.

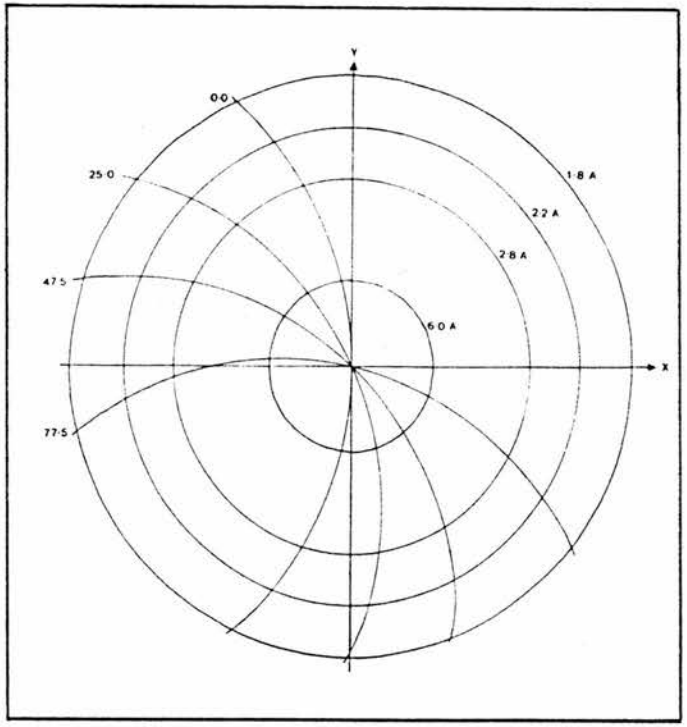
The capillary tube containing the crystal was attached to a goniometer head, and this was then mounted on the oscillation camera in station 9.6 at the SRS, Daresbury. The use of synchrotron radiation in protein crystallography has been well documented [281]. Its excellent collimation produces a very narrow, high intensity x-ray beam that enables higher resolution data to be measured. This was essential for our structure determination, as a previous attempt to record data above 2.5A resolution using a conventional laboratory x-ray source had failed.

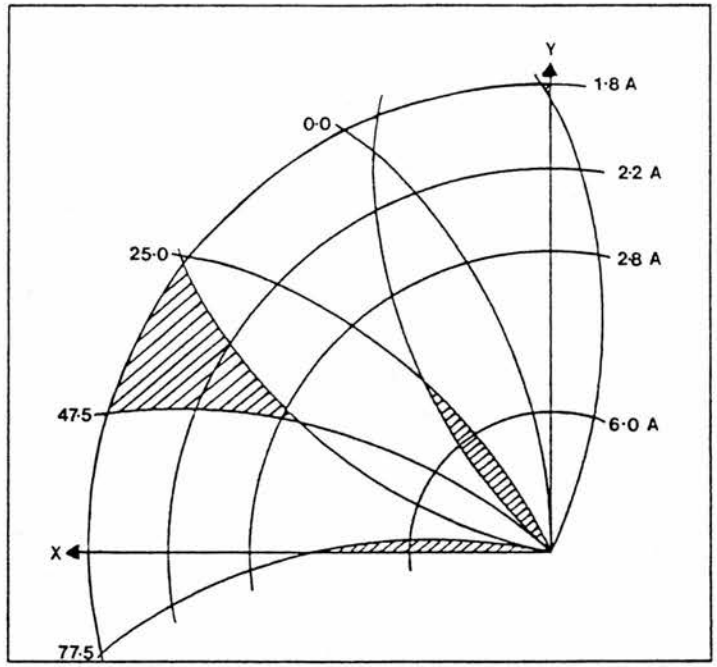
The orientation of the crystal was adjusted until it was aligned, with one of its real lattice vectors coaxial to the camera rotation axis. Still photographs were then

Figure 2:5

DATA REDUCTION GENERATED BY MMM SYMMETRY.

The Ewald sphere.





If data are collected from 0^{o} -> 25^{o} and 47.5^{o} -> 77.5^{o} then the shaded areas represent the uncollected data.

taken at rotation angles of 0^{o} and 90^{o}. In the case of BLG lattice Y each of the other crystal axes, in turn, was positioned along the x-ray beam due to the orthogonal nature of the crystals. From these stills, any misalignment of the crystal axes relative to those of the camera coordinate system could be determined.

Data were then collected on five native orthorhombic BLG lattice Y crystals ($B22_12$, a=55.7A, b=67.2A, c=81.7A) to 1.8A resolution. The SRS x-ray beam had been focussed using a platinum-coated mirror, and passed through a singly bent triangular germanium crystal monochromator before interacting with the crystal [281]. The diffracted radiation was then detected on an Arndt-Wonacott rotation camera [280]. The reflections were recorded on CEA Reflex 25 flat photographic film. Each of the eight cassettes on the octahedral carousel contained three pieces of film(A, B and C with A nearest the crystal) which were separated from one another by 200μm thick sheets of aluminium foil. Marks(fiducials) had previously been placed on three corners of each film, using a low-powered x-ray beam, to enable the subsequent determination of the film orientation relative to that of the camera. The carousel, spindle rotation and exposure times (100-250s/^{o}) were all controlled by microprocessors, thus limiting the amount of manual intervention required. The films were developed immediately by putting them into developer (12.4l water + 2.6l Kodak LX24) for 4.5 minutes, stop (14.4l water + 0.6l Kodak indicator stop) for 30 seconds and then fixer (11.6l water + 3l Kodak FX40 + 375ml Hardener HX-40) for 10 minutes. They were then washed with water for at least 30 minutes and dried in a hot air cabinet at 60^{o}C. A summary of the data collection conditions used for each crystal is given in Figure 2:6.

Figure 2:6

DATA COLLECTION CONDITIONS FOR FIVE NATIVE BLG LATTICE Y CRYSTALS.

	Crystals				
	A	**B**	**C**	**D**	**E**
Spindle axis	b^*	b^*	a^*	a^*	b^*
Oscillation range(°)	2.0/2.5	2.5	2.5	1.25	2.0
Rotation range(°) : from	0	35	357.5	339	355
to	39.5	57.5	32.5	1.5	17
Number of packs	18	9	14	18	11
Wavelength(A)	1.003	1.003	0.874	0.874	0.874
Crystal-to-film(mm)	87	87	110	110	110

2:3 DATA PROCESSING.

Data processing for BLG lattice Y was carried out on a Vax 11/750 using the CCP4 suite of programs supplied by the SRS facility at Daresbury. They are summarized in Figure 2:7.

2:3:1 DIGITIZATION OF THE FILMS.

The recorded intensities were digitized by scanning the films on a Joyce-Loebl Scandig 3 microdensitometer linked to a PDP11/45 computer at Leeds University. (Each film was fixed on a rotating transparent cylindrical drum, and the transmission of light through it was measured). A film area of 120mm x 120mm was scanned in 50μm raster intervals and the intensities recorded on 2400ft magnetic tape at 1600bpi, prior to downloading on to a VAX 11/750. The relationship between the orientation in which the films were scanned, and that in which the data were collected, is important for processing, and is illustrated in Figure 2:8.

2:3:2 DETERMINATION OF THE ORIENTATION MATRIX.

The ideal reciprocal lattice coordinate system is related to the orthogonal camera coordinate system (X_c, Y_c, Z_c) by the orientation matrix, [A]. However this matrix does not allow for any misalignment of the crystal which often occurs in practice. To correct for this, the transformation between the camera coordinates and the Miller indices is given by

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = [A].\phi_x.\phi_y.\phi_z. \begin{pmatrix} h \\ k \\ l \end{pmatrix}$$

where ϕ_x, ϕ_y and ϕ_z are the missetting angles. These angles were obtained from the two still photographs

Figure 2:7

FLOW CHART SUMMARIZING THE CCP4 PROGRAMS USED IN ROTATION CAMERA FILM PROCESSING.

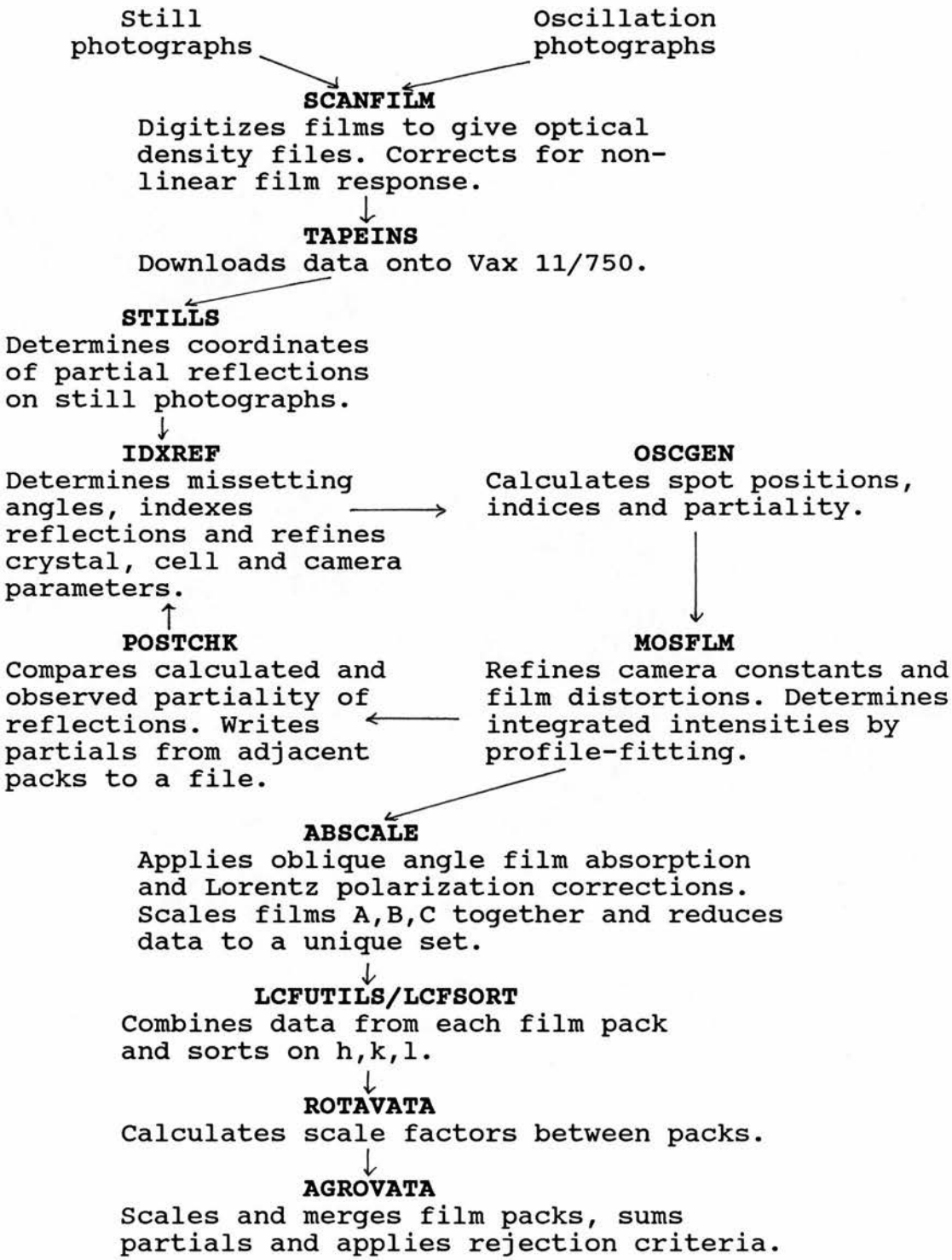
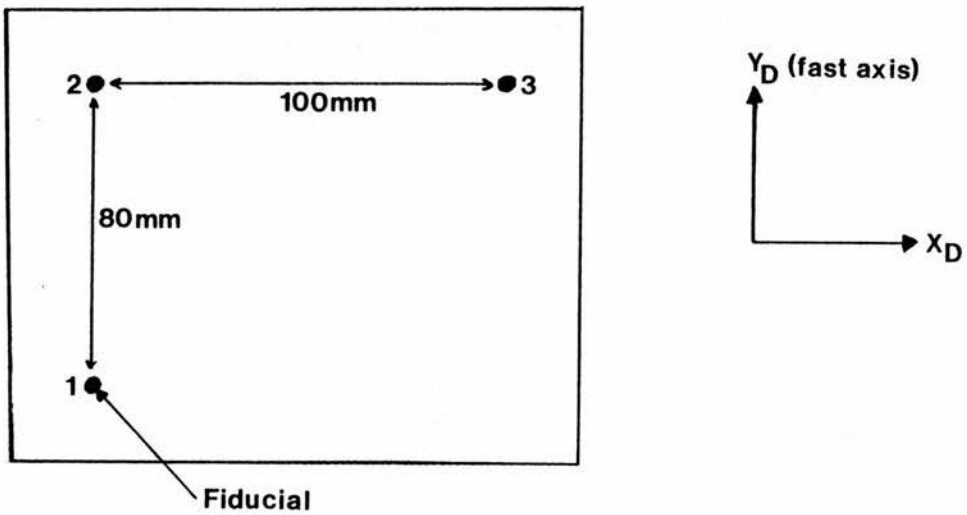


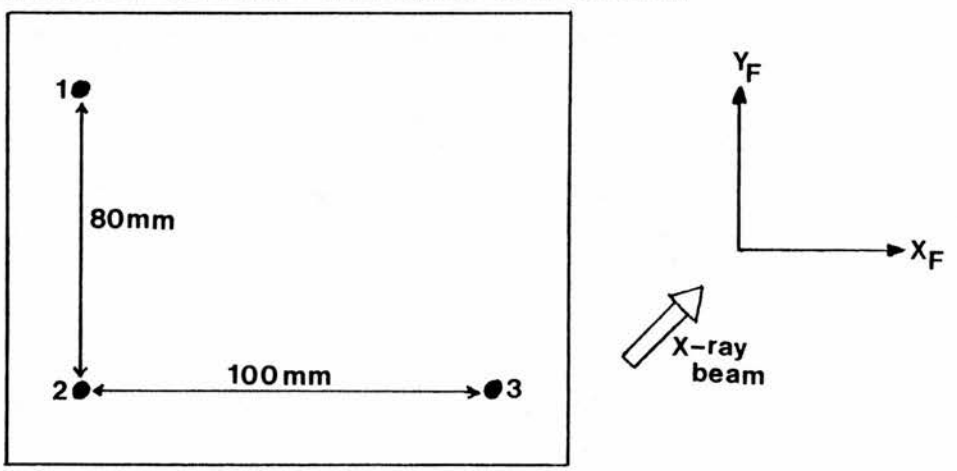
Figure 2:8

GEOMETRICAL RELATIONSHIP BETWEEN THE ORIENTATION OF THE FILM FOR DATA COLLECTION AND SCANNING.

Orientation of film on drum for scanning



Orientation of film for data collection



$$\therefore \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X_D \\ Y_D \end{pmatrix} = \begin{pmatrix} X_F \\ Y_F \end{pmatrix}$$

recorded at 0^{o} and 90^{o} for each of the five native BLG lattice Y crystals.

The reflections on these stills were then indexed, their Miller indices being forced to be integral, as Bragg reflections must lie on the Ewald sphere. Cell parameters, missetting angles, and the crystal-to-film distance were then refined by minimizing the function

$$\sum_{i=1}^{N} ((Ro_i - Rc_i)/ d^*)^2$$

where Ro_i and Rc_i are the observed and calculated distances of the i^{th} reflection from the origin of reciprocal space. The refined cell parameters obtained for each crystal are given in Figure 2:9.

2:3:3 OBTAINING INTEGRATED INTENSITIES.

The integrated intensity for each recorded reflection was determined by profile-fitting [285,286].

The initial step was the location of the fiducials: weak fiducials were detected by lowering the measurement threshold, whereas the few fiducials missed during data collection were either generated from the other films in the pack, or by aligning films from consecutive packs.

The observed diffraction pattern was then compared with the calculated coordinates of the expected reflections. The failure of spots to lie within 80μm of their expected positions is often due to crystal slippage during data collection, and can be remedied by updating the orientation matrix and missetting angles. The differences between the observed and calculated positions of the central low order reflections were then refined by a least squares(LS) procedure to obtain accurate values for the film distortions. Optical densities for all the spots on the film were then recorded using a measurement box consisting of a central region containing the whole

Figure 2:9

REFINED VALUES OF THE CELL PARAMETERS, MISSETTING ANGLES AND CRYSTAL-TO-FILM DISTANCE USED AT THE BEGINNING OF DATA COLLECTION.

		Crystals				
		A	**B**	**C**	**D**	**E**
Missetting angles(°):	ϕ_x	-0.405	-0.916	0.046	0.766	0.617
	ϕ_y	0.659	-0.632	0.183	0.131	-0.120
	ϕ_z	0.220	-0.500	-2.391	0.270	-0.135
Cell parameters:	a	55.70^	55.53	55.70^	55.70^	55.70^
	b	67.20^	67.00	67.20^	67.20^	67.20^
	c	81.70^	81.46	81.70^	81.70^	81.70^
Crystal-to-film distance(mm)		87.004	87.000^	110.646	110.203	110.525

Key

^ : parameter unrefined

peak and a surrounding region of background not encroached upon by adjacent peaks.

The final stage of the BLG data processing was to obtain the integrated intensities by applying profile-fitting, as this helps eliminate random errors and improve the R_{merge} values for datasets containing a large number of weak reflections [285]. The film area was divided into thirteen boxes and a standard profile generated for each of them from the optical densities observed after subtraction of the background. If the profile failed to meet the rejection criteria (eg. made up of less than 10 reflections or with too high a rms background) adjacent profiles were averaged. In the extreme cases where averaging failed to produce an acceptable profile, a 'top-hat' profile was used and yielded the simple integrated intensity. The profiles obtained could be accumulated from more than one A-film, before their application to the B and C films.

The BLG data from all five crystals were processed in this manner, and typical values obtained at different stages of this procedure are given in Figure 2:10.

2:3:4 CORRECTIONS TO THE DATA.

A variety of correction factors should be applied to the data at different stages of processing.

The non-linear film response to x-ray radiation indicated that measured optical densities(ODm) were not linearly related to the intensities falling on the film, and that a parabolic correction was required [289]. The corrected optical densities(ODc) for the BLG data were calculated from

$$ODc = ODm + \{ODm^2 / (2 \times ODmax)\}$$

during the scanning stage of data processing (Chapter 2:3:1).

Figure 2:10

TYPICAL VALUES OBTAINED DURING THE DETERMINATION OF THE INTEGRATED INTENSITIES FROM A SINGLE FILM PACK.

Film A

For superposition of calculated and observed patterns:
rotation of calculated pattern = -0.23^{o}
shift in the centre = -5.0, 77.3 (in 10μm units)
Two LS refinement cycles over low order reflections:
number of spots used = 29
starting residual = 15.0
final residual = 1.9
One cycle of LS refinement over the entire film:
number of spots = 60
starting residual = 3.4
final residual = 2.5
Number of reflections:
number of reflections = 1262
number of full reflections = 1026
number of overloads = 195
Profiles obtained from thirteen regions of the film:
number of acceptable profiles = 9
number of profiles obtained by averaging = 4
number of 'top-hat' profiles = 0
R-factors: $R_{full} = 7.7$
$R_{partial} = 17.5$
Number of bad spots = 1

Film B

Two cycles of LS refinement over the entire film:
number of spots = 38
starting residual = 4.8
final residual = 2.9
Number of reflections:
number of reflections = 557
number of full reflections = 500
number of overloaded reflections = 100
R-factors: $R_{full} = 5.1$
$R_{partial} = 9.0$
Number of bad spots = 1

Film C

Two LS cycles of refinement over the entire film:
number of spots = 25
starting residual = 3.6
final residual = 1.8
Number of reflections:
number of reflections = 355
number of full reflections = 314
number of overloaded reflections = 44
R-factors: $R_{full} = 4.8$
$R_{partial} = 8.6$
Number of bad spots = 0

An oblique angle of incidence correction was necessary to compensate for the differential absorption of the x-ray beam by the films within a pack, at various angles of incident radiation [280]. The integrated intensities obtained by profile-fitting were corrected according to the equation:

$$I_C = I_O \cdot \{e^{-a-(N-1)B} - e^{-a-NB}\} / \{e^{(-a-(N-1)B)\sec 2\theta} - e^{(-a-NB)\sec 2\theta}\}$$

where a and B are constants that depend upon the path length and incidence area of the x-ray beam at the film, and N is the number of films in a pack [213].

The Lorentz(kinematic) factor, L, is applied to reflections to correct for their varying transit times through the Ewald sphere. As it is proportional to $(\sin 2\theta \sin\alpha)^{-1}$ (the angles are defined in Figure 2:4) the intensities close to the rotation axis are unreliable [280].

The polarization factor, PZ, is used to correct for the polarization of the x-ray beam as it passes through the monochromator [290]. For a rotation camera and synchrotron radiation the polarization correction is

$$PZ = C\cos^2 2\theta + D$$

where C and D are constants that depend upon the parallel and perpendicular components of polarization, and the monochromator crystal [281]. Both corrections were applied to the BLG intensities after the films within a pack had been scaled together (Chapter 2:3:5).

A crystal absorption correction may also be needed to allow for the pathlength travelled by the beam through the crystal at various settings. As this correction is difficult to calculate, it is better to use near-spherical crystals and neglect this effect. The lattice Y crystals of BLG used had dimensions of 0.5mm x 0.3mm x 0.2mm. A correction for the absorbance of the capillary tube was not required due to the high transmittance of the Lindemann tube.

2:3:5 DATA REDUCTION.

Data reduction is achieved by scaling together the films within a pack, then the packs for any individual crystal, and then the data from the individual crystals. The scale factors used for all these scalings were calculated from

$$K_i = K_i' \exp\{(-2B_i \sin^2\theta) / \lambda^2\}$$

and then refined using the scaling routine of Fox and Holmes [292]. The quantity

$$\overline{\{I(hkl) - [I(hkl)/K_i]\}^2 / sigI(hkl)^2}$$

was minimized by an iterative LS refinement procedure.

The scale factors between the three films in a pack were generated by taking the ratios of the mean intensities as the initial scale factors. These were then refined by minimizing the above quantity using three cycles of LS refinement, and applied to the data. The scaled measurements were averaged to reduce the data to a unique set. The mean scale factors used and the R_{scale} obtained for each BLG crystal are given in Figure 2:11.

Once the films within each pack had been scaled together, all the data for any individual crystal were combined, sorted, and scaled together. The initial scale factors were obtained from the ratios of the mean intensities between overlapping packs for any one crystal, and then both the scale and temperature factors were refined using the Fox and Holmes scaling routine [292], until convergence was achieved ie. both parameters gave shift/SD < 0.2. The scale factors were then applied to BLG data from that crystal, partially recorded reflections were summed, and equivalent reflections - whether due to repeated measurements or symmetry equivalents - were averaged. Analyses of the data present for each crystal are summarized in Figures 2:12, 2:13 and 2:14.

Figure 2:11

THE SCALING OF FILMS, WITHIN EACH PACK, FOR BOVINE BLG CRYSTALS.

	Crystals				
	A	**B**	**C**	**D**	**E**
Number of packs	18	9	14	18	11
Total rotation range collected (°)	39.5	22.5	35.0	22.5	22.0
Reflections measured:					
total number	28362	15162	20166	14233	13438
those with multiplicity=1	22472	11446	17713	10448	10977
Mean scale factors:					
film A to B	2.24	2.48	2.70	2.70	2.68
film A to C	5.23	5.83	7.20	7.32	7.18
Mean R_{scale} to 1.8A (%)	3.1	3.5	5.3	3.5	3.1

Figure 2:12

ANALYSIS OF SCALED DATA FROM EACH BOVINE BLG CRYSTAL.

	Crystals				
	A	**B**	**C**	**D**	**E**
Total number of reflections	22006	10450	17524	10446	10928
Mean intensity	407	1665	642	945	634
Mean SD	0.155	0.201	0.388	0.117	0.275
Number of full reflections	20596	10049	16918	9942	10656
Number of partial reflections	1410	401	606	504	272
Fractional bias of partials	-0.0217	-0.0259	-0.1084	-0.0734	-0.0919
Independent hkl	8026	5107	5841	4204	3948
% > 3 SD(I) (to 1.8A resolution)	65.8	76.1	64.5	71.0	69.1
R_{merge} on I (to 1.8A resolution)	0.076	0.062	0.108	0.047	0.067

Figure 2:13

VARIATION IN R_F AS A FUNCTION OF RESOLUTION FOR DATA FROM EACH BOVINE BLG CRYSTAL.

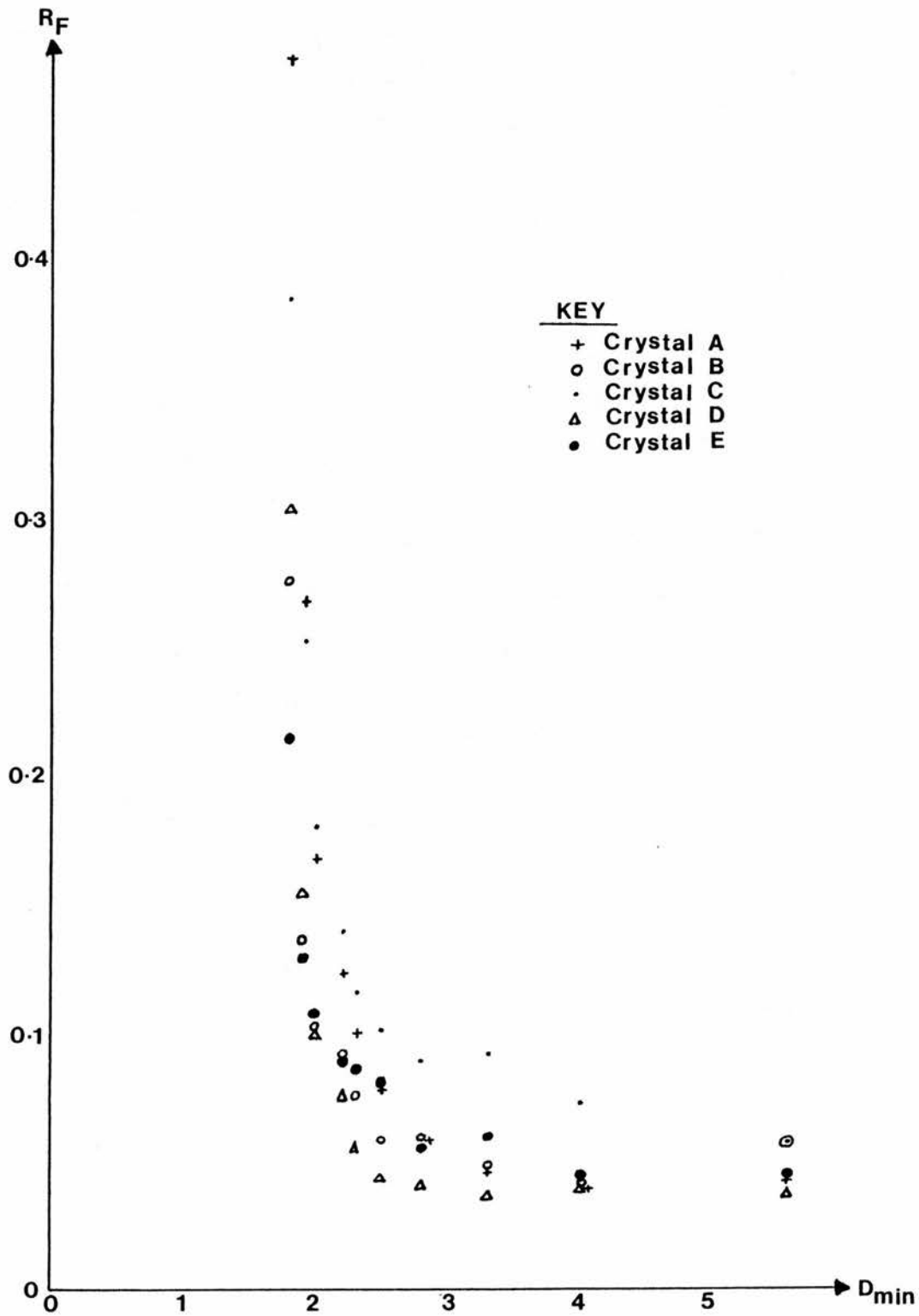
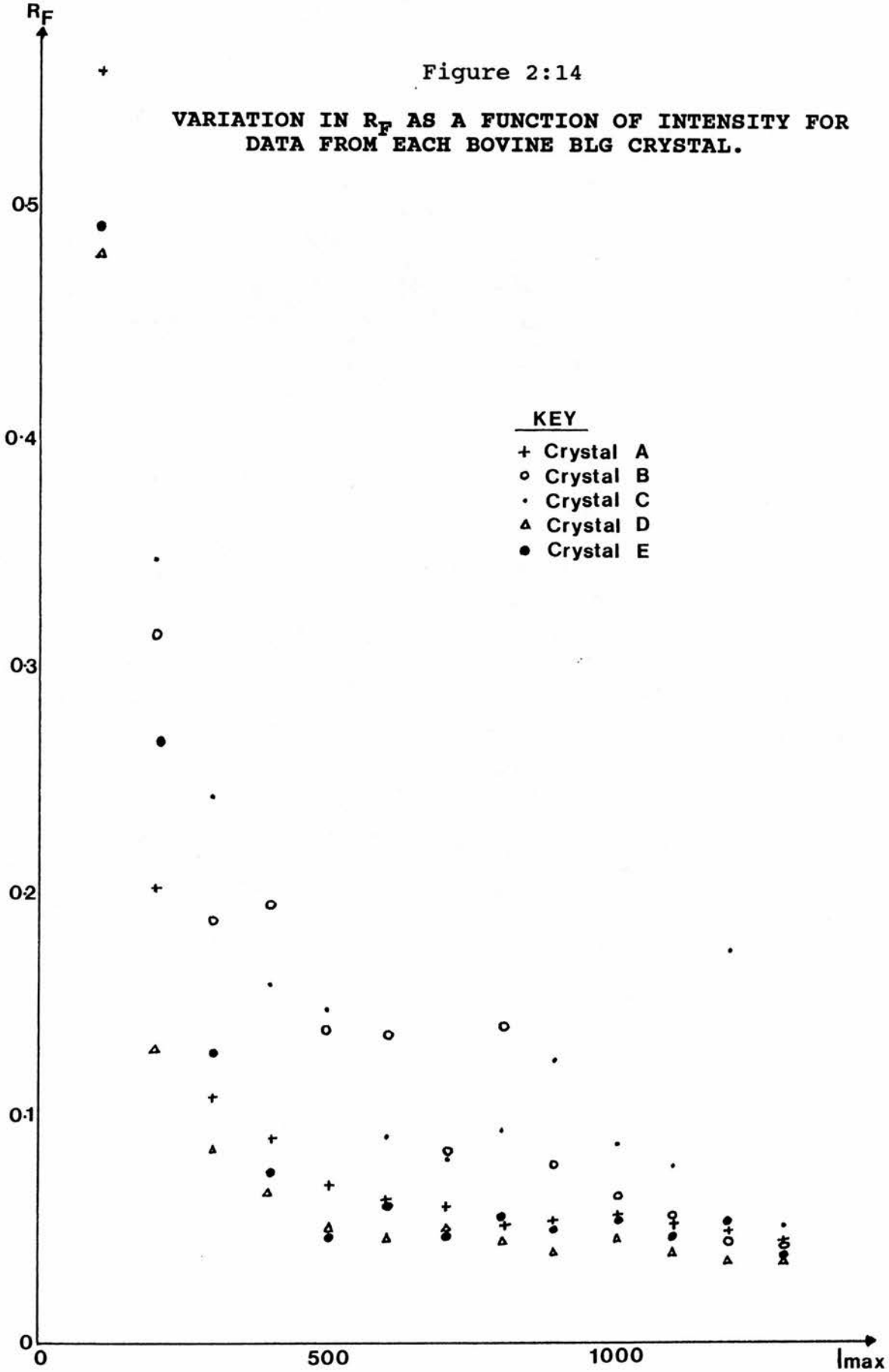


Figure 2:14

VARIATION IN R_F AS A FUNCTION OF INTENSITY FOR DATA FROM EACH BOVINE BLG CRYSTAL.



The data collected on each of the crystals were then merged together, to give a unique set of data containing h, k, l, F, and sigF values for each reflection. A summary of this dataset, as a function of resolution, is presented in Figure 2:15. It was noted that the distribution of the scaled reflections (as measured by rms sig_{inp}) did not correlate well with that of the observed data prior to scaling (as measured by sig_{obs}) (Figure 2:16). Thus the constants SDFAC and SDADD in the equation

$$(\text{rms } sig_{obs})^2 = \{\text{SDFAC} \times (\text{rms } sig_{inp})^2\} + (\text{SDADD} \times I^2)$$

were adjusted to reduce the discrepancy between distributions of observed and scaled data. The data to 1.8A resolution was then rescaled, and a much better fit between sig_{obs} and sig_{inp} was noticed, although the R_F to 1.8A had deteriorated to 10.8%. The five packs of data with R_F > 14% were then rejected and the data rescaled using the scale and temperature factors obtained by the Fox and Holmes method [292] and given in Figure 2:17. The resultant dataset was then analysed {Figures 2:18 and 2:19} and the percentage of unique reflections as a function of resolution shells is given in Figure 2:20. This revealed that the dataset contained most of the reflections up to 1.9A resolution, and that the missing data were predominantly the low-to-medium resolution reflections.

2:4 ADDITION OF LOW RESOLUTION DATA.

Several sets of low and medium resolution data had been collected on lattice Y crystals of bovine BLG by previous workers. These sets were now combined with the 1.8A dataset. The programs used for this combination, and

Figure 2:15

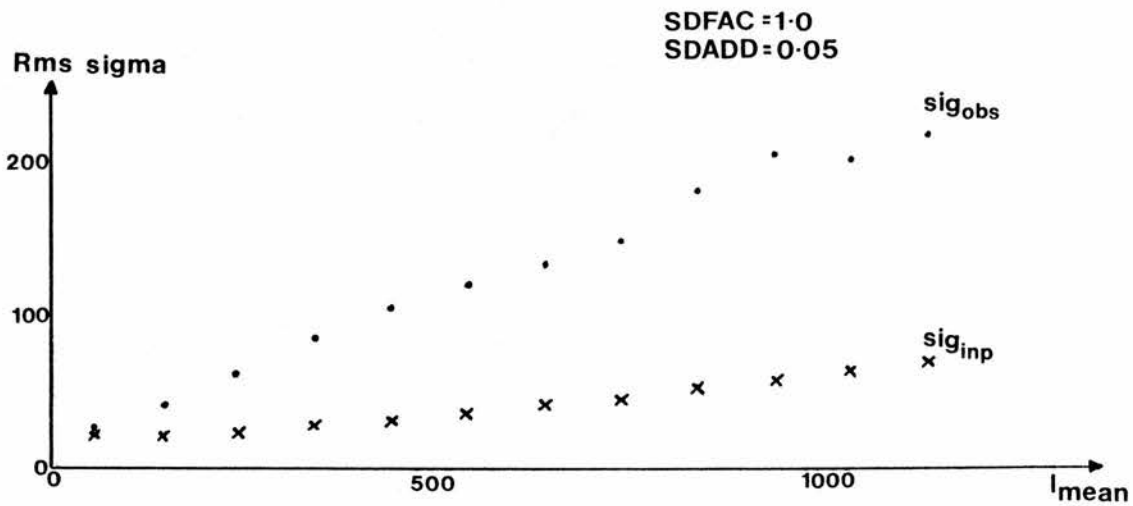
SUMMARY OF THE PROCESSED DATA FROM BOVINE BLG LATTICE Y COLLECTED ON THE SYNCHROTRON.

	Resolution (in A)				
	3.0	**2.5**	**2.2**	**2.0**	**1.8**
No. measurements	13321	25306	38210	51856	68644
No. reflections	2700	4954	7458	10086	13762
Independent hkl	2109	3757	5436	7065	8742
% unique dataset	78.5	84.5	87.6	89.5	89.6
% > 3SD(I)	97.2	95.1	92.7	88.6	78.5
R_F (%)	8.2	8.4	8.6	9.0	9.7

Figure 2:16

VARIATION IN THE DISTRIBUTION OF THE OBSERVED AND SCALED (INPUT) REFLECTIONS AS A FUNCTION OF INTENSITY.

Plots indicate the discrepancy between distributions using the default values of the constants SDFAC and SDADD, and the match achieved after adjusting the constants.



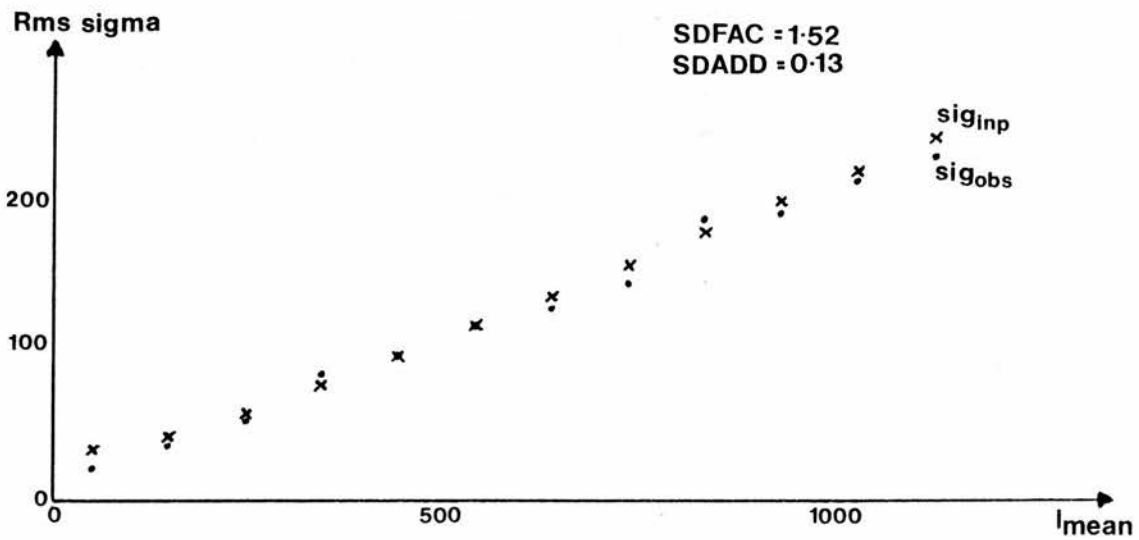


Figure 2:17

THE SCALE AND TEMPERATURE FACTORS USED IN THE FINAL SCALING OF THE FIVE NATIVE BLG LATTICE Y CRYSTALS.

Convergence achieved after 5 cycles: shift(K) / SD(K) = 0.047
shift(B) / SD(B) = 0.074

Residual = 2.79

PACK	K	SD(K)	B	SD(B)	PACK	K	SD(K)	B	SD(B)	PACK	K	SD(K)	B	SD(B)
1001	1.0000	0.0407	0.000	0.523	1002	1.1166	0.0412	-0.524	0.474	1003	1.1880	0.0436	-0.872	0.475
1004	0.8096	0.0349	-2.958	0.533	1005	0.8699	0.0358	-2.179	0.498	1006	0.9656	0.0380	-2.529	0.494
1007	0.8943	0.0362	-4.400	0.496	1008	0.9726	0.0388	-4.567	0.517	1009	1.0209	0.0403	-5.607	0.524
1010	1.1030	0.0468	-5.814	0.574	1011	0.2220	0.0108	-2.949	0.530	1014	0.6239	0.0248	-2.865	0.489
1016	0.6358	0.0295	-4.244	0.593	1017	0.5619	0.0249	-6.607	0.549	2001	0.2464	0.0107	-1.005	0.472
2002	0.3149	0.0130	-1.530	0.447	2003	0.4261	0.0177	-2.795	0.488	2004	0.4295	0.0200	-1.078	0.524
2005	0.3717	0.0158	-2.691	0.473	2006	0.3762	0.0160	-3.432	0.488	2007	0.4059	0.0180	-3.686	0.525
2008	0.3767	0.0183	-4.819	0.571	2009	0.2921	0.0138	-4.100	0.539	3001	0.7222	0.0326	0.292	0.563
3002	0.3463	0.0164	-1.068	0.546	3003	0.4513	0.0198	-1.688	0.530	3004	0.4540	0.0194	-4.676	0.526
3005	0.5705	0.0282	-2.525	0.605	3006	0.5424	0.0226	-3.883	0.523	3007	0.6017	0.0270	-3.845	0.552
3008	0.5472	0.0233	-6.369	0.543	3010	0.4753	0.0211	-4.903	0.532	3011	0.4374	0.0200	-7.488	0.563
3012	0.5843	0.0273	-4.314	0.615	3013	0.5461	0.0246	-6.221	0.567	3014	0.5738	0.0290	-8.010	0.650
4001	0.3801	0.0281	-1.836	0.862	4002	0.3772	0.0249	-2.820	0.819	4003	0.3429	0.0246	-3.916	0.851
4004	0.3306	0.0227	-3.624	0.899	4005	0.3754	0.0305	-1.869	0.951	4006	0.3643	0.0270	-2.826	0.940
4007	0.3376	0.0267	-4.082	0.916	4008	0.3369	0.0234	-5.050	1.096	4009	0.4107	0.0276	-4.490	0.818
4010	0.4075	0.0347	-5.352	1.021	4011	0.3791	0.0334	-5.568	0.973	4012	0.3694	0.0299	-6.188	1.000
4013	0.3551	0.0276	-5.697	0.984	4014	0.3692	0.0337	-7.067	0.907	4015	0.4194	0.0307	-6.262	0.939
4016	0.3762	0.0259	-7.611	0.868	4017	0.3224	0.0281	-5.682	1.059	4018	0.3054	0.0209	-7.274	0.847
5001	0.7870	0.0342	-1.093	0.537	5002	0.7473	0.0329	-1.476	0.498	5003	0.7046	0.0323	-1.051	0.577
5004	0.4692	0.0200	-1.502	0.510	5005	0.4938	0.0219	-1.956	0.540	5006	0.5136	0.0251	-3.047	0.586
5007	0.5353	0.0245	-3.476	0.544	5008	0.5168	0.0232	-4.636	0.533	5009	0.4647	0.0208	-5.346	0.535
5010	0.4646	0.0202	-5.723	0.543	5011	0.4886	0.0227	-5.513	0.554					

Figure 2:18

ANALYSIS OF THE RESULTANT DATASET, TO 1.8A RESOLUTION, COLLECTED FROM BOVINE BLG LATTICE Y.

Number of full measurements used : 63873
Mean intensity : 465
Mean SD (as a function of intensity) : 0.249

Number of partials used : 0

Number of reflections : 12825
Unique dataset contains 15360 reflections
therefore % unique dataset present : 83.5%

R-factor : 0.097 to 1.8A resolution

% > 3 SD(I) : 71.2%

$(I(J)-I_{mean})/SD(J)$	No.	
-4.5	3	
-4.0	6	
-3.5	13	
-3.0	15	
-2.5	208	*
-2.0	307	*
-1.5	626	***
-1.0	1990	**********
-0.5	6802	**********
0.0	15445	**********
0.5	21409	**********
1.0	9771	**********
1.5	2916	**********
2.0	910	****
2.5	270	*
3.0	151	
3.5	44	
4.0	28	
4.5	49	
5.0	10	
5.5	0	

Figure 2:19

VARIATION IN R_F AND '% REFLECTIONS > 3SD(I)', AS A FUNCTION OF RESOLUTION, FOR THE BLG LATTICE Y DATA COLLECTED ON THE SYNCHROTRON.

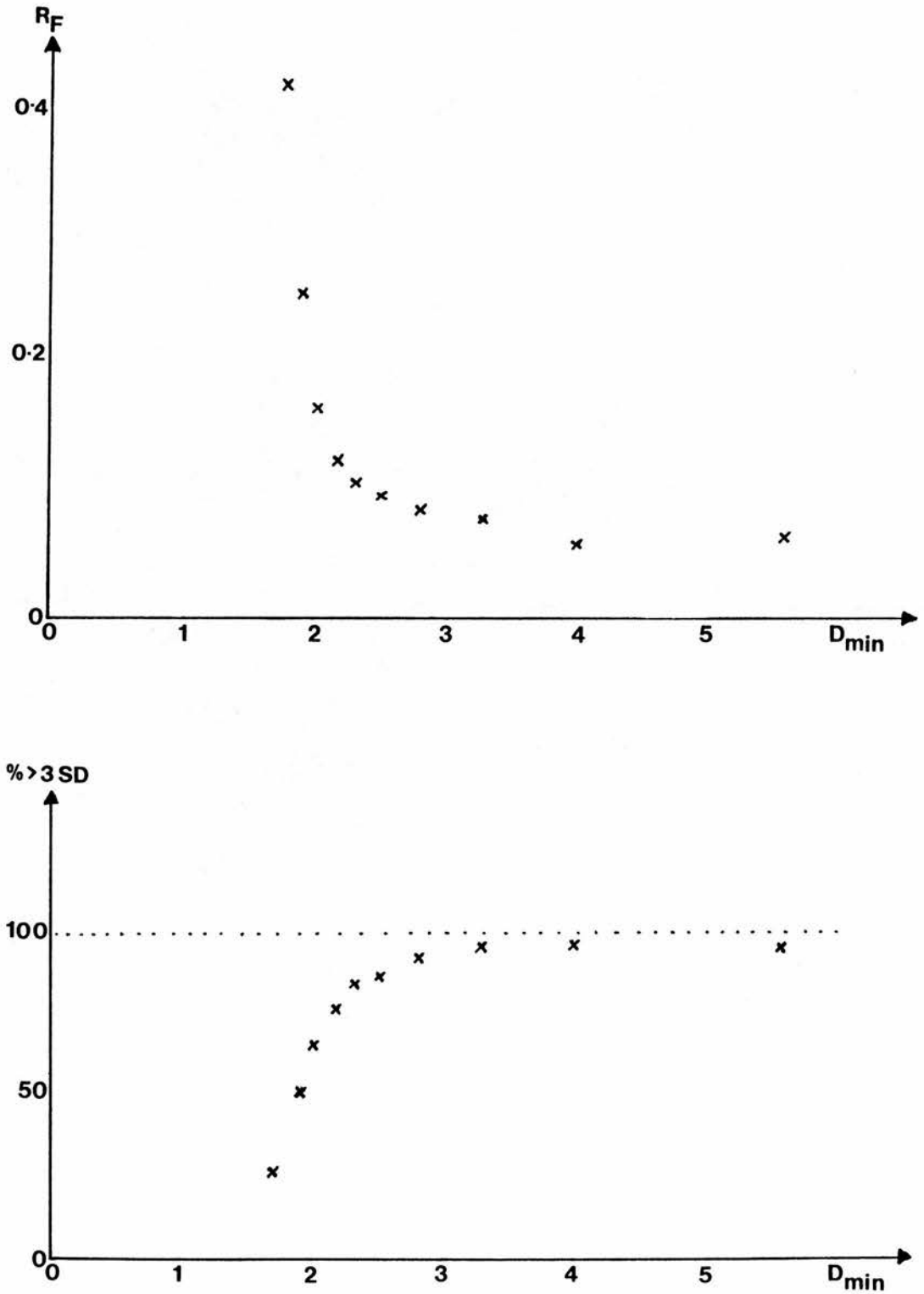


Figure 2:20

THE PERCENTAGE OF THE UNIQUE DATASET PRESENT AT VARIOUS RESOLUTIONS.

Resolution (A)	% Unique dataset
5.0	59.0
3.6	75.3
2.9	91.6
2.5	92.6
2.3	93.9
2.1	94.6
1.9	95.0
1.8	67.0

the subsequent steps involved in the generation of an electron density map, are outlined in Figure 2:21.

2:4:1 SCALING THE DATA.

D.W.Green had collected a 3.3A dataset on the linear diffractometer at the Royal Institution, but the reflections recorded did not have standard deviations associated with them. Their $sigF_D$ values were estimated as

$$sigF_D = F_D^{0.5} + 0.5$$

and then appended to the data.

M.Z.Papiz collected two sets of medium resolution data using copper Kα radiation. One was collected to 2.8A resolution using an Enraf-Nonius CAD4 diffractometer; whereas the other was collected on an Arndt-Wonacott rotation camera to 2.2A resolution, although the reflections between 2.5A and 2.2A were relatively weak [213].

The structure factors and standard deviations from the above three datasets: F_D and $sigF_D$ from the 3.3A data, F_4 and $sigF_4$ from the 2.8A CAD4 data, and F_R and $sigF_R$ from the 2.2A rotation dataset, were all concatenated to those F and sigF present for reflections in the 1.8A resolution dataset.

The squares of the structure factors F, F_R and F_D were each compared to those of F_4, and a scaling function used to determine the overall scale factors. For each reflection F, F_R and F_D were then individually scaled to F_4. The scale factors used, and the R_{scale} values obtained between each set of scaled structure factors, are given in Figure 2:22. Although the R_{scale} values appear poor - possibly due to the data having been collected on instruments with differing geometries, the applied scale factors being unrefined, and no temperature factor being included - the constancy of the ratios F_4/F,

Figure 2:21

PROGRAMS USED TO COMBINE BLG LATTICE Y DATASETS AND GENERATE AN ELECTRON DENSITY MAP.

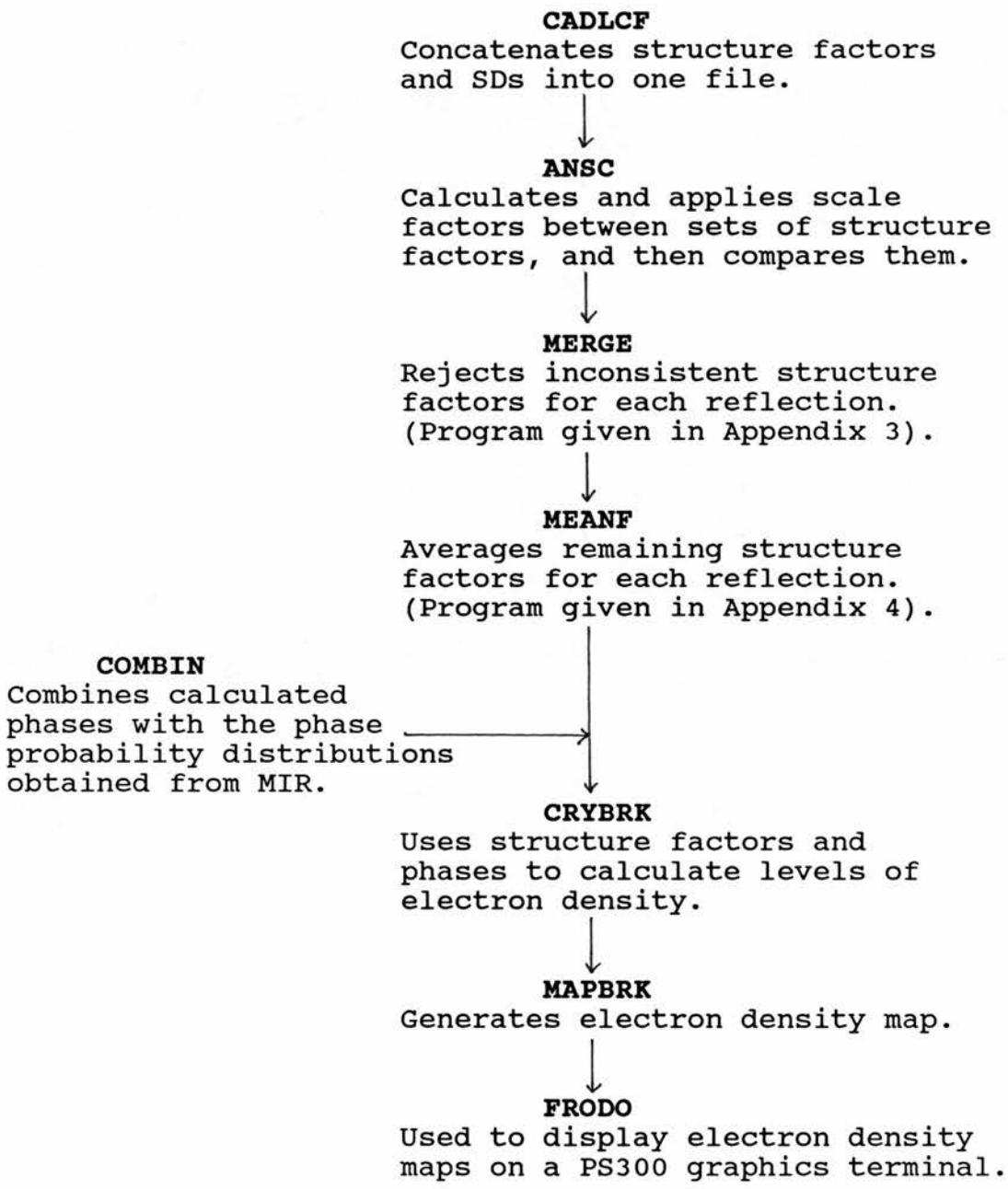


Figure 2:22

ANALYSIS OF THE SCALING TOGETHER OF FOUR BLG LATTICE Y DATASETS.

Structure factors to be scaled together	Number of common reflections	Scale factors (on F)	R_{scale} (on F)
F to F_4	2245	3.64	23.6%
F_R to F_4	2403	3.66	14.8%
F_D to F_4	805	3.62	13.2%

F_4/F_R and F_4/F_D as a function of resolution suggested that the scaled data was reasonable {Figure 2:23}.

2:4:2 MERGING THE DATA.

The dataset for bovine BLG lattice Y now contained up to four structure factor measurements per reflection. Thus it seemed sensible to reject any inconsistent measurements, and then average the remaining values. A program was written in Fortran77 {Appendix 3} to implement the rejection criteria given in Figure 2:24, and any inconsistent reflections were then rejected.

All the remaining structure factors for each reflection were averaged together using the program MEANF {Appendix 4}. For each reflection, F and sigF were first converted to I and sigI by

$$F = I^{0.5}$$

$$sigF = -(I)^{0.5} + (I+sigI)^{0.5}$$

and then scaled together on I. The data were subsequently converted back to F and sigF values. This produced a dataset containing 13944 reflections, 90.8% of the unique data to 1.8A resolution. The distribution of these reflections with resolution is shown in Figure 2:25. This revealed that the low and medium resolution shells had been virtually filled by the addition of the other three datasets, and that the resultant dataset was now nearly complete to 1.9A resolution.

2:5 GENERATION OF AN ELECTRON DENSITY MAP.

The electron density, $\rho(xyz)$, at the position (xyz) within a unit cell, is given by

$$\rho(xyz) = [1/V]\Sigma F(hkl)\ e^{i\alpha(hkl)} e^{-2\pi i(hx+ky+lz)}$$

The structure factor amplitude, F(hkl), for each reflection is obtained from the square root of the

Figure 2:23

VARIATION IN THE RATIOS OF THE SCALED STRUCTURE FACTORS AS A FUNCTION OF RESOLUTION.

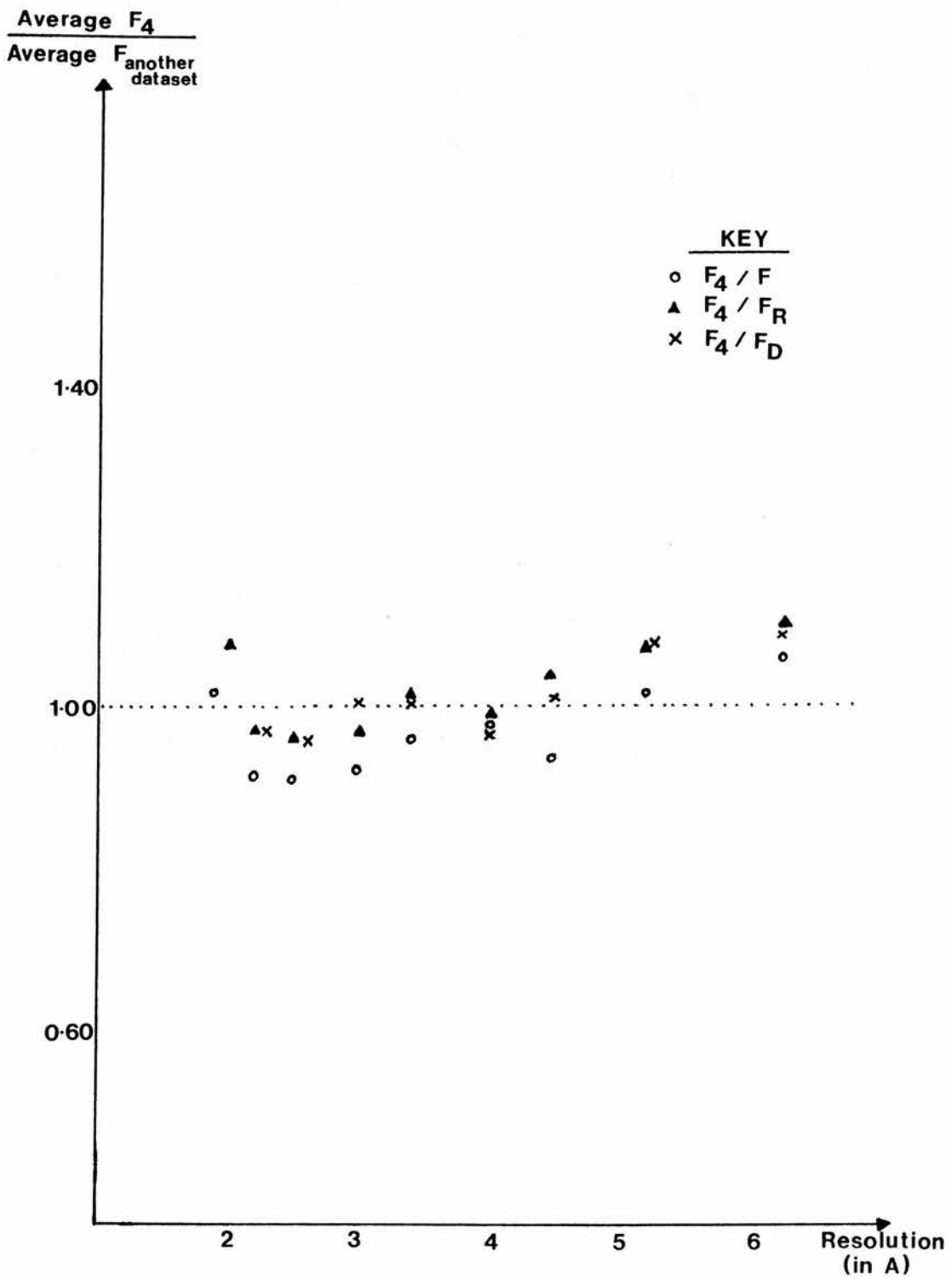


Figure 2:24

CRITERIA FOR REJECTING INCONSISTENT REFLECTIONS.

1: If only one reflection is present for a specific (hkl) then it is kept.

2: If all the reflections for any specific (hkl) obey
$$(F_{mean} - F(J)) / SD(J) < 3$$
then they are all kept.

3: If two reflections are present and
(a) one fails the above criteria, then it is set to zero, and the other reflection is kept.
(b) both fail the above criteria, the smaller one is set to zero, and the other kept.

4: If three or four reflections are present, and one or more fail the above criteria, the one which has the largest value of
$$(F_{mean} - F(J)) / SD(J)$$
is set to zero.

Figure 2:25

PERCENTAGE OF THE UNIQUE DATASET PRESENT, AS A FUNCTION OF RESOLUTION, IN THE RESULTANT BOVINE BLG DATASET.

Resolution (A)	% Unique dataset
5.0	99.9
2.9	99.7
2.5	93.5
2.3	93.9
2.1	94.6
1.9	95.0
1.8	67.0

intensity recorded on the diffraction pattern, but the phase angle, $\alpha(hkl)$, is more difficult to determine. From a knowledge of these two quantities, for each reflection, an electron density map can be calculated which represents the structure of the crystal.

2:5:1 METHODS OF DETERMINING PHASE ANGLES.

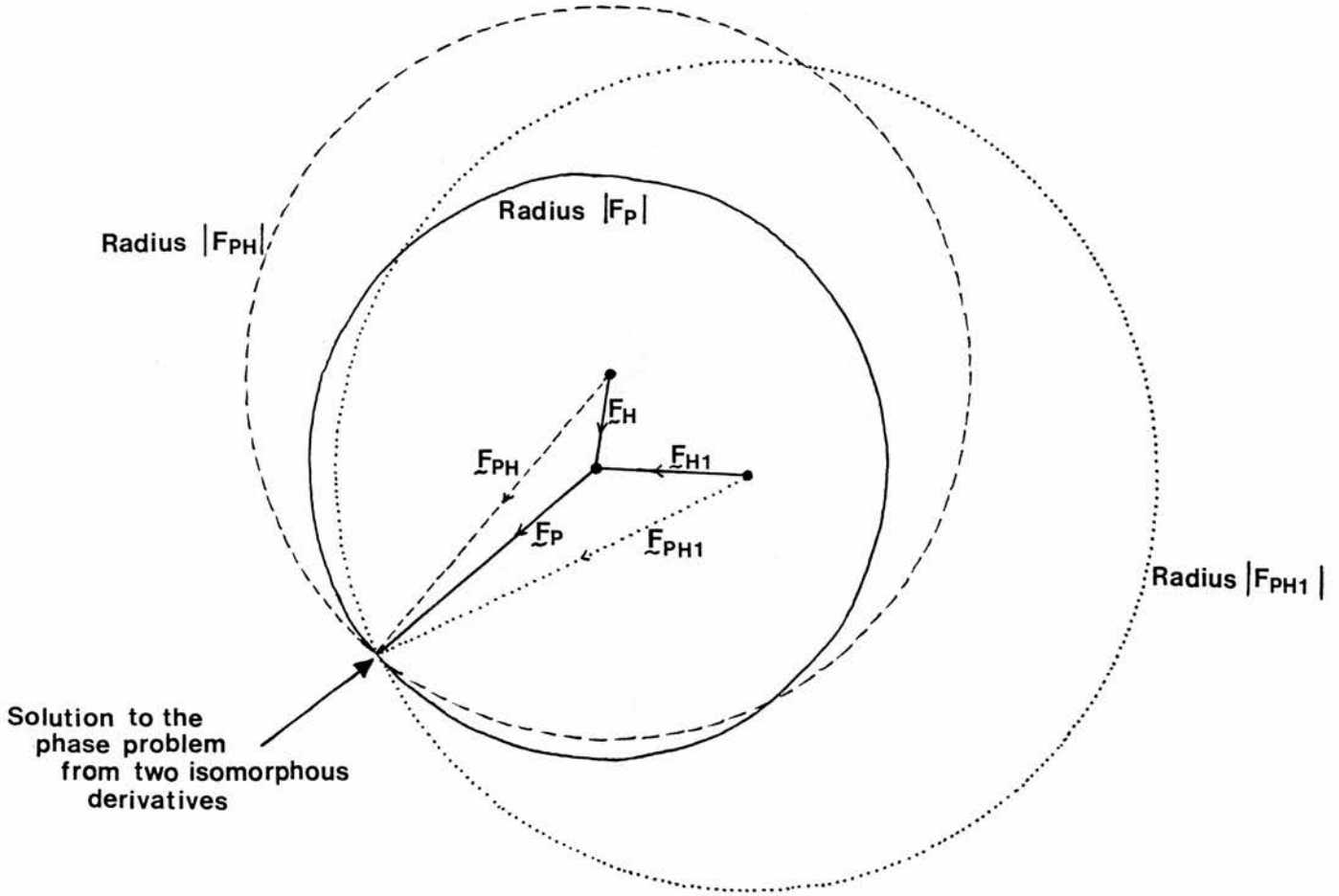
In protein crystallography phase angles can be determined by multiple isomorphous replacement(MIR), multiwavelength anomalous dispersion, or molecular replacement.

The method of MIR involves the preparation of heavy atom derivatives of the native crystals. The data collected on these derivatives must show isomorphism with the native data, as any perturbation to the molecules or crystal lattice invalidates the condition $\mathbf{F}_{PH} - \mathbf{F}_{P} = \mathbf{F}_{H}$. For each derivative the position of the heavy atom is then determined by either direct methods [294] or a difference Patterson synthesis. Its coordinates, together with the scattering factors of the heavy atoms, are then used to calculate the structure factors for the heavy atom contribution. This is done for at least two derivatives, and then the phases can be determined vectorially from a Harker construction which is illustrated in Figure 2:26.

Anomalous scattering is a small effect that arises from a heavy atom within the molecule, and leads to the breakdown of the Friedel Law between symmetry-related reflections [296]. Recent developments in the tunability of the SRS now allow anomalous dispersion data to be recorded at several wavelengths near then absorption edge of the heavy atom. The differences in the anomalous scattering at the various wavelengths, and between Friedel pairs, enable the phases to be calculated in a manner analogous to that used in MIR.

Figure 2:26

A HARKER CONSTRUCTION SHOWING THE VECTORIAL DETERMINATION OF PHASES FROM TWO ISOMORPHOUS DERIVATIVES.



Molecular replacement is a technique which enables phase information from a known structure to be applied to an unknown one. Self-Patterson functions for both structures are generated, and summed over all rotations of one relative to the other, so as to determine the rotation function. Structure factors are then calculated from the known structure, as a function of position, and the R_F values used to obtain the translation function [295]. The known structure is then placed in the same orientation as the unknown one, and phases are calculated which approximate to those for the real structure.

2:5:2 DETERMINATION OF PHASES FOR BLG.

The phases associated with the low-to-medium resolution reflections recorded from BLG lattice Y crystals had been determined previously [213]. Data had been collected on four heavy atom derivatives: up to 2.5A resolution from the PTN derivative, and up to 2.2A resolution for the MMA, HgI_2, and HgI_4^- derivatives. The heavy atom positions were determined from a combination of the low resolution work [5] and the calculation of difference Patterson syntheses with coefficients $(F_{PH}-F_P)^2$. The single site occupied in the MMA derivative, the major site in the HgI_4^- derivative, and the major and minor sites in the PTN derivative were used independently to calculate initial phases to 2.8A resolution for each dataset. The phased sets were then combined, and the phase coefficients A, B, C and D for each reflection summed. The phase probability distributions were then determined from

$$P(\alpha) = \exp (A\cos\alpha + B\sin\alpha + C\cos 2\alpha + D\sin 2\alpha)$$

[299]. These distributions were subsequently improved by the inclusion of calculated phase information from the partial structure using Sim weighting [298,300], and by solvent flattening.

2:5:3 GENERATION OF AN ELECTRON DENSITY MAP.

The phases generated above were appended to the observed structure factors of each reflection recorded up to 2.8A resolution, and then a Fourier synthesis calculated. The initial 2.8A electron density map generated by M.Z.Papiz was from a Fourier synthesis using coefficients $mF_{obs}\exp(i\alpha)$ where the figure of merit was 0.703 [213]. After the addition of our data, but still using the same phases, another 2.8A electron density map was calculated. This map had sampling intervals of 64, 80, 96 along the a, b, c axes respectively, and was in sections perpendicular to the c axis. The portion of the map produced was

$$-10/64 < x < 45/64$$
$$-24/80 < y < 48/80$$
$$-30/96 < z < 40/96$$

and this was displayed on the Evans & Sutherland PS300 graphics system using the programme FRODO [301]. A typical region of the map is illustrated in Figure 2:27.

2:6 REFINEMENT.

2:6:1 METHODS OF REFINEMENT.

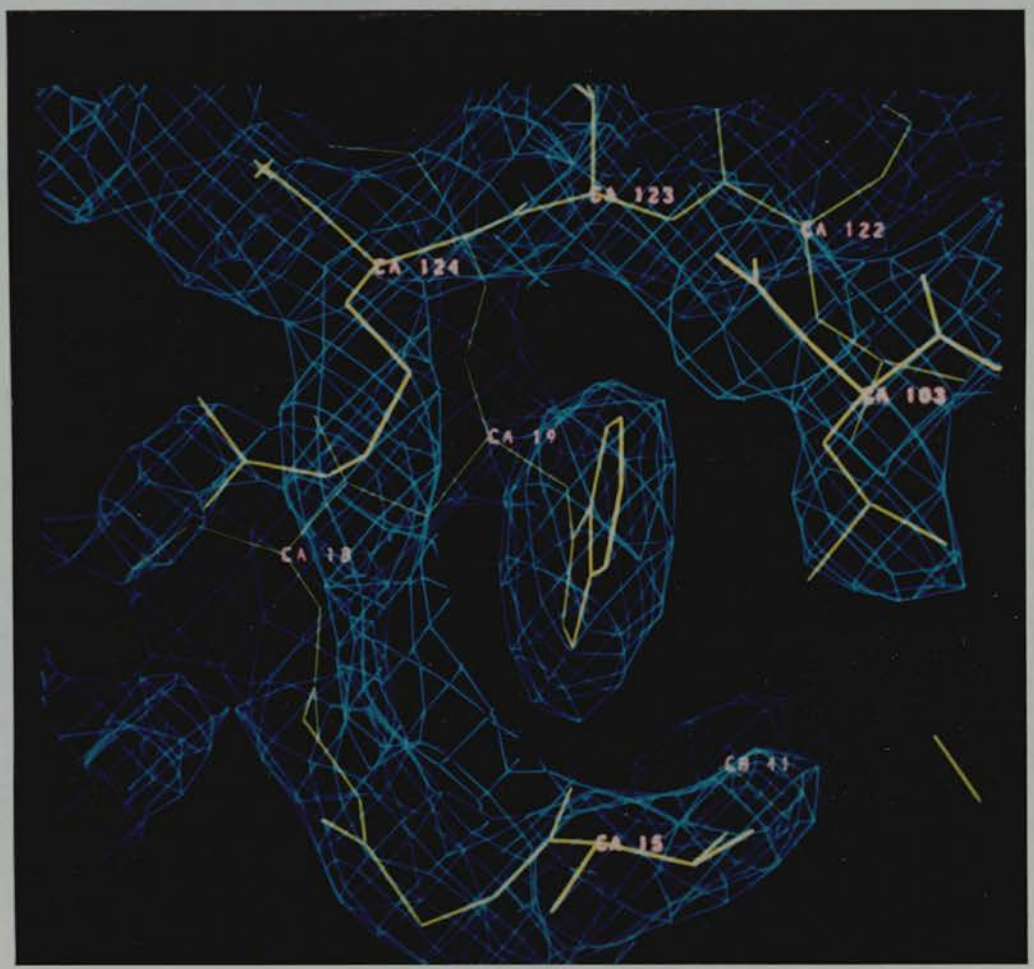
The final stage of a structure determination is the refinement of the atomic and thermal parameters. This can be achieved by several methods: real space refinement, reciprocal space refinement and molecular dynamics refinement.

Real space refinement is a cyclic process involving density modifications which produce a gradual improvement in the phases [302]. The current model is used to calculate ρ_{calc}, and then various physical constraints eg. positivity, atomicity or solvent flattening are

Figure 2:27

A REPRESENTATIVE PORTION OF THE 2.8A ELECTRON DENSITY MAP FOR BOVINE BLG LATTICE Y.

The region shown illustrates the fit of Trp-19 to the electron density.



applied, and the function $(\rho_{obs}-\rho_{calc})^2$ minimized by a least squares procedure [303]. The improved model is then used to recalculate phases.

Reciprocal space refinement methods have now superseded the real space technique. The Hendrickson-Konnert method [305] minimizes the residual

$$\Sigma\ \omega.(F_{obs} - F_{calc})^2 + \Sigma\ \omega.Va^2$$

where Va is the variance of the interatomic distance distribution, and depends upon both the stereochemistry and temperature factors of the atoms. In contrast, Tronrud et al. [306] minimize the function

$$\Sigma\ \omega.(kF_{obs} - F_{calc})^2$$

the thermal parameters being included in the overall 'mismatch', as no exponential term is applied to F_{calc}. Stereochemical constraints are introduced to improve the low ratio of observations to parameters, and then LS minimization is carried out using first-order derivatives and fast Fourier transform algorithms [307]. The initial cycle of refinement determines the scale factor, by minimizing

$$\Sigma\ \omega.\{kF_{obs}-\exp(-B\sin^2\theta/4\lambda)F_{calc}\}^2$$

and this is then used in subsequent cycles of refinement.

Molecular dynamics is incorporated into crystallographic refinement by the addition of an extra 'x-ray energy' term to its conventional empirical potential energy(PE) calculation.

$$\text{New PE} = (1/sig^2)\ \Sigma\ (kF_{calc}-F_{obs})^2 + \Sigma\ (\text{conventional PE terms})$$

The resultant forces are then used to solve Newtons equations of motion, and provide the new acceleration, velocity and position for each atom [308]. By repeating this step many times a large number of conformations are explored, as the kinetic energy of the system enables energy barriers to be crossed. This method has the ability to move atoms out of false local minima and hence a higher radius of convergence, more than 3A compared to

1A for LS refinement. It also gives better R_F values than LS refinement alone, but is not as good as manual model building combined with LS refinement [309]. However manual model building is a time-consuming process, but often necessary to help correct for the large errors in side-chain positions or the misassignments of amino acids. Difference Fourier syntheses with coefficients $(F_{obs}-F_{calc})$ or $(2F_{obs}-F_{calc})$ are used to generate electron density maps which can then be displayed on a graphics terminal. Atoms, residues or substructures are then superimposed, and can be moved individually and built into the density of the map.

2:6:2 REFINEMENT OF BLG.

A variety of methods have been used so far in the refinement of the atomic and thermal parameters of BLG lattice Y: manual model building, restrained LS refinement, and molecular dynamics.

The 2.8A resolution electron density map generated from our resultant dataset appeared superior to that obtained by M.Z.Papiz [279], and had a R_F of 48%. The BLG polypeptide was superimposed upon this map by using the programme FRODO [301] on an Evans & Sutherland PS300 graphics system. Both individual atoms and amino acid residues were manually adjusted, and geometrical constraints applied regularly using the REFI option. As there appeared to be no clear main chain electron density for residues Glu-55 -> Lys-69 (Figure 2:28) or Ser-150 -> Ile-162, these regions were initially left unaltered. Different approaches to the refinement of this initial model-built structure were then tried, and these are summarized in Figure 2:29.

The coordinates were sent to J. Brady at Cornell University who kindly carried out MD simulations on them. The shifts observed in the atomic positions seemed

Figure 2:28

PHOTOGRAPH TO ILLUSTRATE THE LACK OF ELECTRON DENSITY FOR THE MAIN CHAIN BETWEEN RESIDUES GLU-55 AND LYS-69.

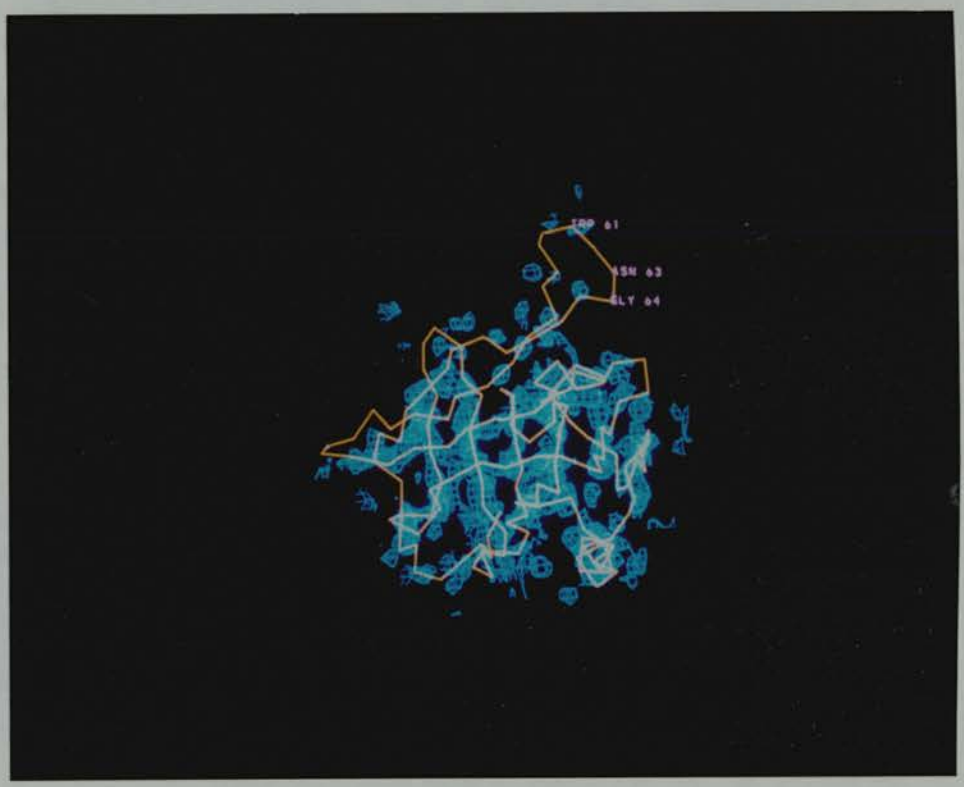
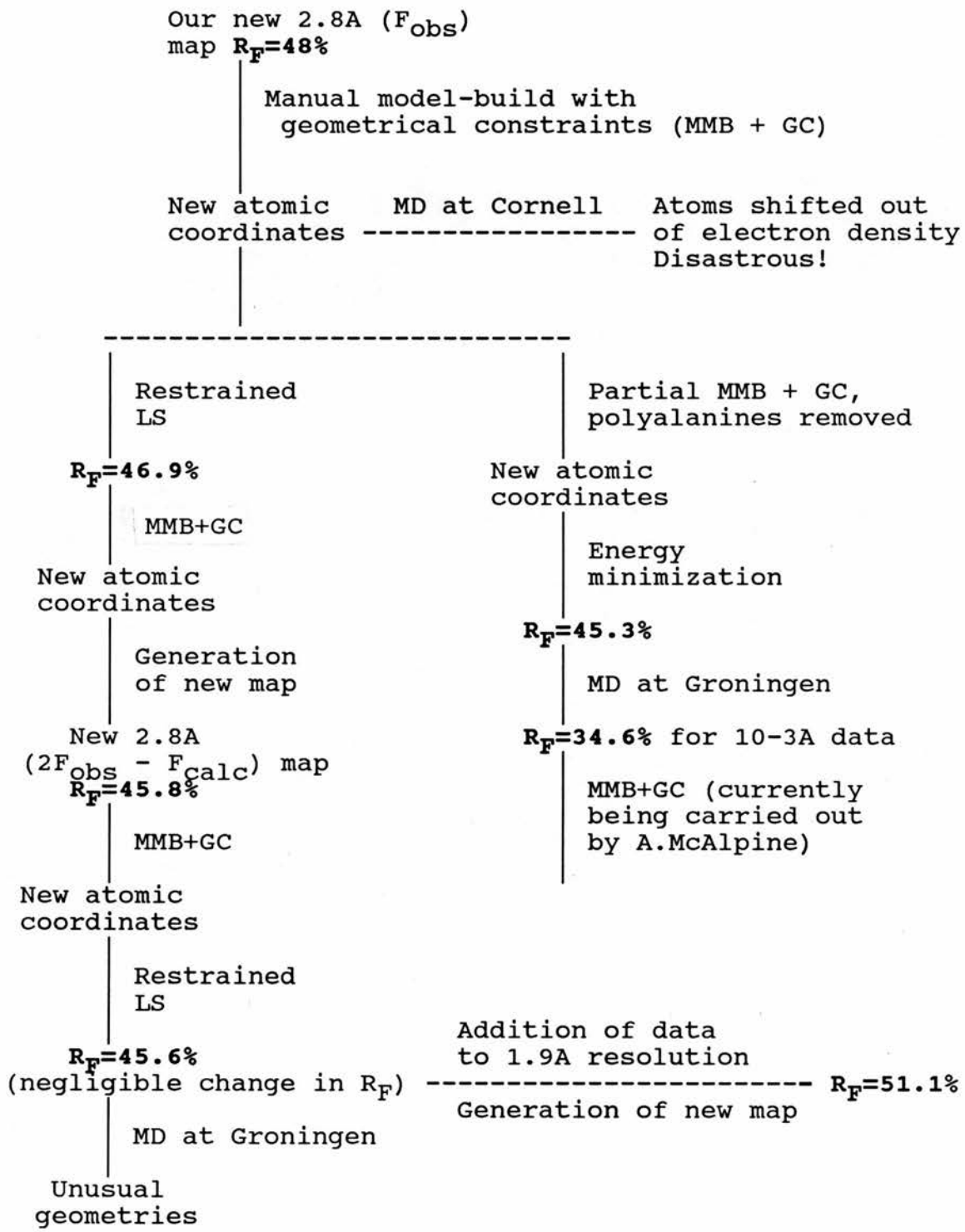


Figure 2:29

OVERVIEW OF THE REFINEMENT OF THE BLG LATTICE Y STRUCTURE.



disastrous as the molecule had been pulled away from the electron density. This was not surprising as no x-ray constraints had been applied and the primary sequence was incomplete due to the presence of some polyalanine residues. Thus these coordinates were disregarded.

Restrained LS refinement was then applied to the model-built structure, using the programme TNT [306] - and three cycles of refinement were required for convergence. The R_F decreased to 46.9% and a significant adjustment to the position of the loop 55-69 was observed, although these residues still failed to correspond to any main chain electron density. Another manual model-build was then carried out, but the location of the flexible loop 55-69 still caused problems. At various stages a few of the residues within this loop were removed, and new maps generated - but these failed to show any significant differences when compared to the initial map. The possibility of using the structure of an homologous protein eg. RBP or alternate crystal forms of BLG to locate this loop were considered but rejected. The loop region was not one of the conserved regions in RBP and had not been located in BLG lattice X. Its position in lattice Z had originated from the initial dubious coordinates used in lattice Y, and thus was not useful.

The resultant coordinates from this model building were used to calculate phases, which were then combined with the initial MIR phases prior to the generation of a new 2.8A $(2F_{obs} - F_{calc})$ map with an R_F value of 45.8% and a figure of merit of 0.806. A third cycle of manual model building with the regular addition of geometrical constraints was carried out using this new map, and then restrained LS refinement was undertaken using a stronger weighting factor for the structure factors than for the geometrical terms. Both the crystallographic and geometric parameters at this stage of refinement are summarized in Figures 2:30 and 2:31. The poor geometric

Figure 2:30

CRYSTALLOGRAPHIC DETAILS OF BLG LATTICE Y OBTAINED AFTER THE FINAL CYCLE OF REFINEMENT.

	Resolution									
	6.83	**5.08**	**4.35**	**3.92**	**3.62**	**3.40**	**3.22**	**3.07**	**2.95**	**2.85**
No. of reflections	415	394	394	388	383	384	392	388	367	381
Weighting of F_{obs} compared to F_{calc}	1.11	1.22	1.30	1.39	1.47	1.55	1.63	1.71	1.78	1.86
R_F	0.48	0.42	0.41	0.47	0.44	0.48	0.44	0.51	0.49	0.51
No. of phased reflections	404	384	389	372	361	367	354	317	281	267
Mean phase shift	-5.10	-3.19	-3.34	-4.37	0.61	-1.53	7.37	12.37	-0.28	3.80
Figure of merit	0.90	0.86	0.85	0.81	0.82	0.79	0.77	0.75	0.74	0.71

Mean R_F is 0.456
Mean figure of merit is 0.806

Figure 2:31

STEREOCHEMICAL DEVIATIONS PRESENT IN BLG LATTICE Y AFTER THE FINAL CYCLE OF REFINEMENT.

	Number of deviations	Mean	RMS
Bond deviations			
Bond lengths(A)	1292	-0.006	0.169
Bond angles(°)	1745	0.702	13.007
Trigonal atom non-planarity(A)	40	-0.021	0.037
Planar groups			
Deviation from plane(A)	176	-0.037	0.058
Torsion angles			
Torsion angles(°)	787	0.071	28.614
Bad contacts			
Bad contacts(A)	576	0.393	0.522
Chiral centres			
Handedness		7 incorrect	

parameters are partially due to the lower weighting used and to the errors still present in the structure. It should be noted, that out of the worst deviations for each type of stereochemical restraint, 70% occur within the ill-defined regions 55-69 and 150-162. Despite the strong weighting of the structure factors, the R_F value failed to decrease significantly upon LS refinement. Inclusion of the data up to 1.9A produced a poorer electron density map with a R_F of 51.1%. This confirmed that the model structure was still incorrect, and suggested that another approach to the refinement was necessary.

The second approach to refinement of this BLG structure, using energy minimization and MD [308], was implemented by L.Sawyer and M.Fujinaga at Groningen University. Energy minimization of the above LS refined structure was unsuccessful because the protein still contained stretches of polyalanines where no side chain density was apparent. The correct sequence was model-built and then fifty steps of energy minimization were applied using the program GROMOS. This energy-minimized dataset was then subjected to MD refinement starting at 300K and using 0.002ps steps. 5A resolution data were used initially, and then the resolution was extended to 3A, and finally to 2A resolution data {Figure 2:32}. The R_F value decreased significantly, being minimal for the 10-3A resolution dataset with a value of 34.6%. A Ramachandran plot of this data was kindly provided by S.J.Yewdall, and revealed that the geometries of many of the amino acids were still far from ideal {Figure 2:33}. Another manual model-build, with geometrical restraints being applied regularly, is currently being carried out by A.McAlpine. This should remove many of the deviations from ideality observed at present, and provide an adequate model from which phase extension can be carried out, and to which higher resolution data can be added.

Figure 2:32

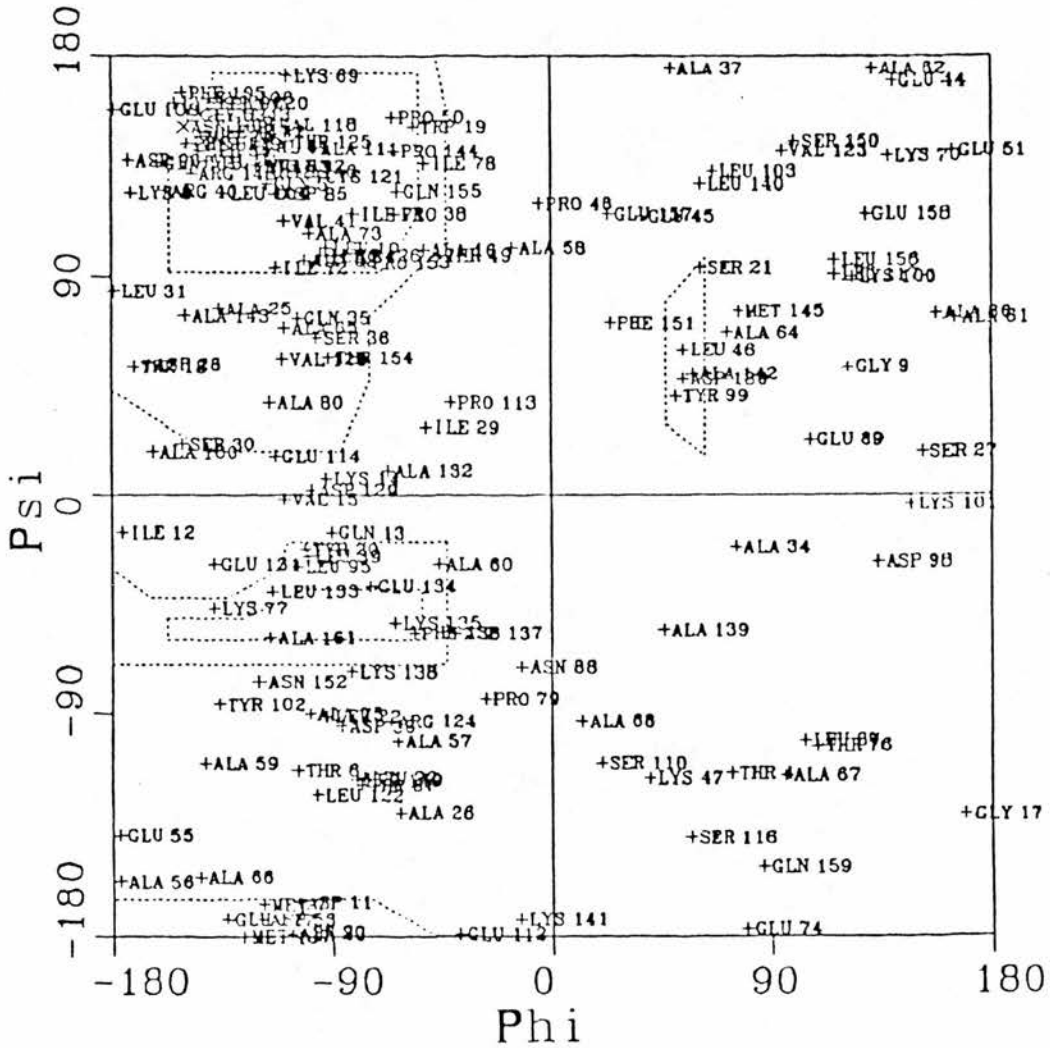
MOLECULAR DYNAMICS REFINEMENT OF BLG LATTICE Y DATA.

The temperature used was 300K with 0.002ps steps.

Resolution range	No. of steps	Weight applied to (F_{obs}-F_{calc})	R_F
10 to 5 A	200	500	44.8%
10 to 5 A	500	300	37.5%
10 to 3 A	200	300	41.5%
10 to 3 A	300	250	37.7%
10 to 3 A	200	200	34.6%
10 to 2 A	100	200	43.2%

Figure 2:33

RAMACHANDRAN PLOT OF BLG LATTICE Y AFTER MOLECULAR DYNAMICS REFINEMENT.



The expected bond angle ranges for regular polypeptide conformations are:

	phi(°)	psi(°)
antiparallel β-sheet	-30 to -170	+40 to +170
α-helix	-60 to -120	-20 to -75

whilst --- indicates the sterically permissible dihedral angles for an alanine residue.

2:7 COMPARISON OF BLG LATTICES Y AND Z.

Crystals of BLG lattices Y and Z have been grown under similar pH conditions. Lattice Y crystals were salted-out from a mixture, the pH of which dropped from 7.8 to 7.3 during their growth. The appearance of many small lattice Z crystals prior to their emergence is consistent with the knowledge that Z crystals can be grown at pH 7.5 [59]. It seems likely that the gradual conversion of lattice Z (pH 7.5) to lattice Y (pH 7.3) involves a slight enhancement in the stability of the lattice Y crystals, relative to that of lattice Z, at the lower pH. However as the core structure of BLG is very similar in both cases [8,59] it is unclear whether the change in stability is purely due to crystal packing effects, or to changes in either the state of association or conformation of the protein over this pH range. Thus a comparison of the two structures was deemed worthy of study.

The state of association of the protein in these crystal forms was investigated. The two-fold symmetry between the positions of heavy atoms in both lattices suggested that the protein was dimeric [304]. Although this is still believed to be the case in lattice Y, the nature of the BLG in lattice Z is less certain - its 2.5A resolution structure suggests that the protein may form a linear polymer, the two contact sites between subunits involving (58-66 and 154-162) and (8-14, 44-48 and 94-99) [59]. These regions are situated on opposite sides of the structure, and neither incorporate Ile-29, His-146 nor Ile-147 which are all part of the dimer interface in lattice Y. This seems to imply that BLG may be monomeric in lattice Z. However the poor quality of the data (keeping reflections with I > 1SD still only gave 63.4% of the unique dataset between 2.7 and 2.5A), the excessive number of parameters refined, and the bad

stereochemistry despite the R_F of 0.22, must throw some doubt on the accuracy of this structure.

The two structures were compared by generating both the average rotation matrix and the translation vector necessary to superimpose the Cα atoms of lattice Z upon those of lattice Y. The matrix:

$$\begin{pmatrix} -0.48249 & 0.47383 & 0.73667 \\ -0.37758 & 0.64637 & -0.66305 \\ -0.79034 & -0.59807 & -0.13296 \end{pmatrix}$$

and vector (17.20393, 18.90215, 66.75434) were then applied to the coordinates of the Cα atoms in lattice Z, and a superposition obtained {Figure 2:34}. The distances, between the positions of the same residues in both structures, after this superposition, are given as a function of residue number in Figure 2:35. Those residues in the β-strands A, B, E, F, G and H (as defined in Figure 1:13) were closer than the rms distance of 5.495A, further supporting the idea that the β-barrel core is similar in both structures. Larger variations were observed between residues in the flexible loop regions {Figure 2:34}; especially 55-69, as expected from the different packing of BLG in the two lattice forms, but also in the adjacent regions. The spatially adjacent loops, 27-34 and 109-117, which lie at the mouth of the cavity were displaced, such that the entrance to the β-barrel appeared more closed in lattice Z. This could explain why, despite the presence of excess retinol, it is only found in the external channel of lattice Z crystals of the BLG+retinol complex [59]. The wider opening in lattice Y should enable retinol to bind in the hydrophobic cavity in a manner analogous to retinol in RBP [215]. The end of the α-helix and strand I were also displaced in lattice Z, so that strand I was no longer H-bonded to the β-barrel - this would alter the interface between subunits and may explain why BLG is monomeric in lattice Z.

Figure 2:34

THE SUPERPOSITION OF THE Cα ATOMS OF LATTICE Z ON THOSE OF LATTICE Y.

Lattice Y is shown in red, whilst lattice Z is in blue.

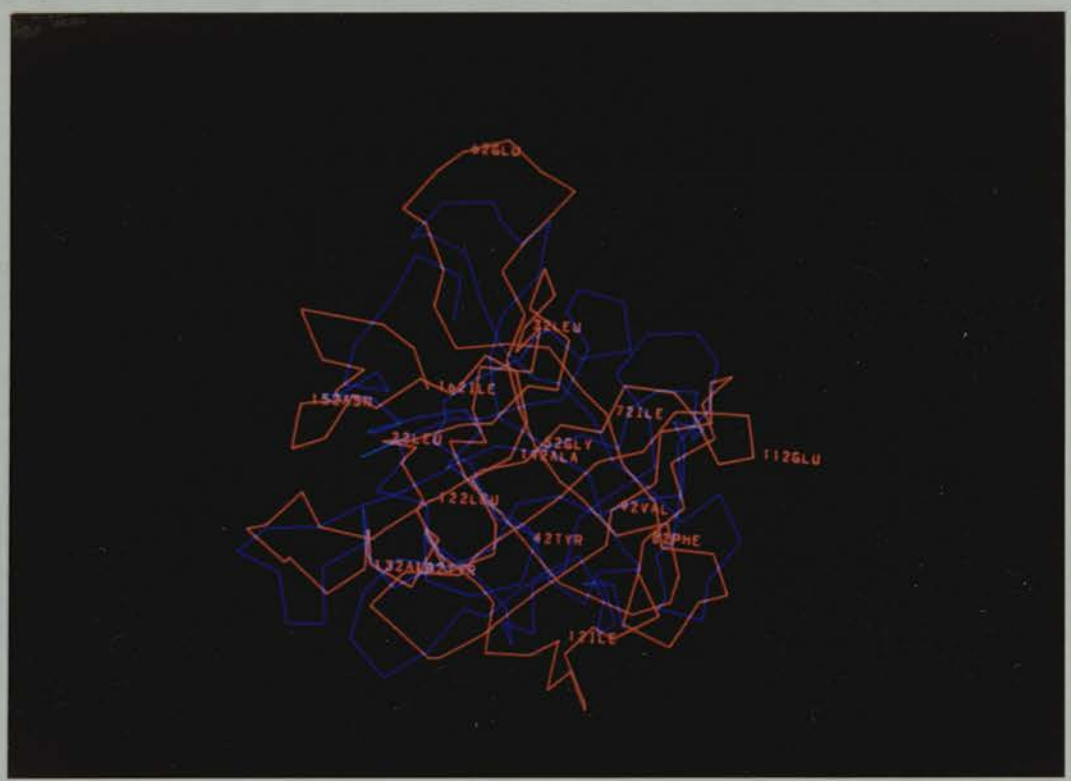
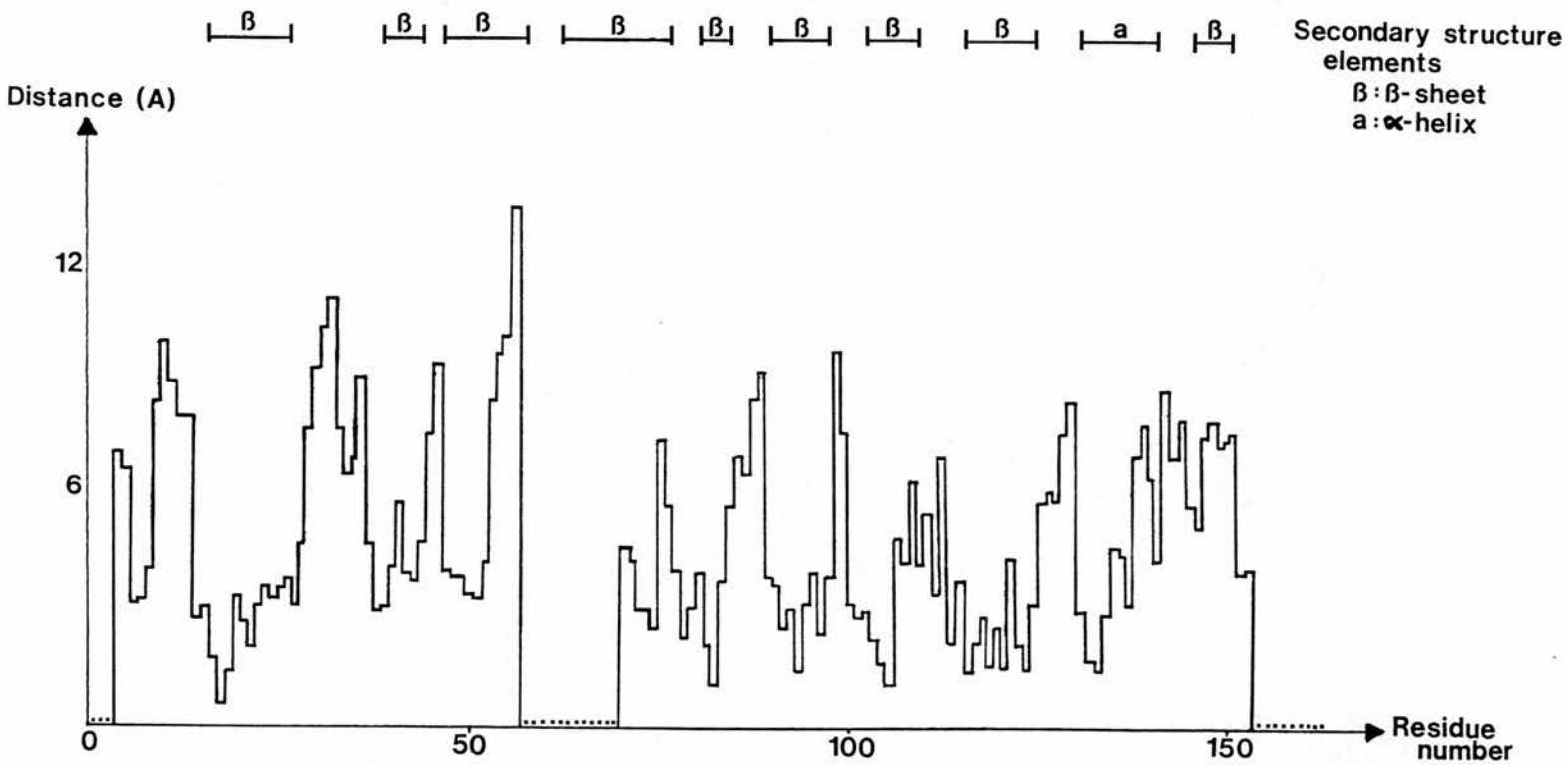


Figure 2:35

AN ANALYSIS OF THE DISTANCES BETWEEN EQUIVALENT RESIDUES IN LATTICES Y AND Z AFTER THEIR SUPERPOSITION.



The observations made fit well with the current knowledge on the binding of retinol to BLG, and suggest that real changes, and not just crystal packing effects have been noticed. However it should be borne in mind that neither structure is known with absolute confidence (due to the poor data used in the lattice Z structure determination, and the ill-defined electron density around 55-69 in lattice Y), and that the superposition is an average. The collection of better data from lattice Z crystals, and extending the resolution of the lattice Y structure using our data, should improve the quality of the comparison between these structures.

2:8 THE DETAILED STRUCTURE OF BLG LATTICE Y.

Prior to this project the current structure of bovine BLG lattice Y was at 2.8A resolution, but unrefined [213]. The addition of our data significantly improved the electron density map, and gave a R_F of 48%. Manual model building and LS refinement then produced a better 2.8A structure whose R_F settled around 45%; whereas energy minimization and MD gave a 3A structure with a R_F of only 34.6%.

2:8:1 THE β-STRUCTURE.

BLG is predominantly a β-protein consisting of nine anti-parallel β-strands and a three-turn α-helix {Figures 1:13 and 2:36}. Of the nine β-strands eight of these (A-H) form two antiparallel β-sheets which wrap around one another, producing a cross-hatched effect and creating a β-barrel. Strand A forms part of both sheets, residues 13-19 combining with those between strands F and G, 97-104, to close the barrel at one end. These amino acids are well conserved in all BLGs, and being located

Figure 2:36

PHOTOGRAPHIC REPRESENTATIONS OF BOVINE BLG LATTICE Y.

The top photograph shows the cross-hatching of the β-strands whilst the bottom one illustrates the β-barrel. The residues labelled are those Cys which form disulphide bridges 66-160 and 106-119, the free thiol Cys-121, Trp-19, and residues 130 and 140 which are situated at the ends of the α-helix.

together [141] {Figure 2:37} may be important for the binding of this protein to a specific membrane receptor [228]. The cavity enclosed has approximate dimensions of (5 x 9 x 8)A^3, Trp-19 at its base, and is lined with hydrophobic residues. Its opening has a rim of hydrophilic amino acids including Lys-70 and Glu-55 which may form a H-bond, Gln-35, and several residues from the regions 85-90 and 108-116 {Figure 2:38}.

The ninth β-strand, strand I (145-150), is believed to be involved in the association of subunits to form a dimer {Figure 2:39}. The hydrophobic interactions between Ile-29, Ile-147 and the packing of symmetry-related His-146 residues are thought to hold the molecules together [8]. However none of these residues are conserved in all dimeric, but absent in all monomeric, BLGs {Figures 1:10 and 1:11}, implying that no single amino acid residue is critical for this association.

Further associations occur within the crystal. A single subunit not only hydrophobically interacts via strand I (145-150) and Ile-29 to form a dimer, but also electrostatically associates with two adjacent molecules {Figure 2:40}. One molecule interacts via the surface containing residues Glu-44, Arg-40, Lys-100 and Lys-124 (the basic nature of the latter is conserved in all BLGs, and may be important for binding BLG to a receptor <u>in vivo</u>); whereas the other molecule interacts via the regions consisting of Lys-70 and Glu-89 (hydrophilic residues located near the entrance of the β-barrel). Both these contacts are due to crystal packing as neither exhibits 2-fold symmetry about the interface.

2:8:2 THE α-HELIX OF BLG.

BLG contains a single, three-turn α-helix (residues 130-140) which is amphipathic in nature {Figure 2:41}. The hydrophilic residues which lie on the outer side of

Figure 2:37

PHOTOGRAPHS TO ILLUSTRATE THE LOCATIONS OF THOSE RESIDUES CONSERVED IN BLGS.

The top photograph shows that these residues are situated near the base of the β-barrel, whilst the bottom one illustrates their grouping in a small area of the protein.

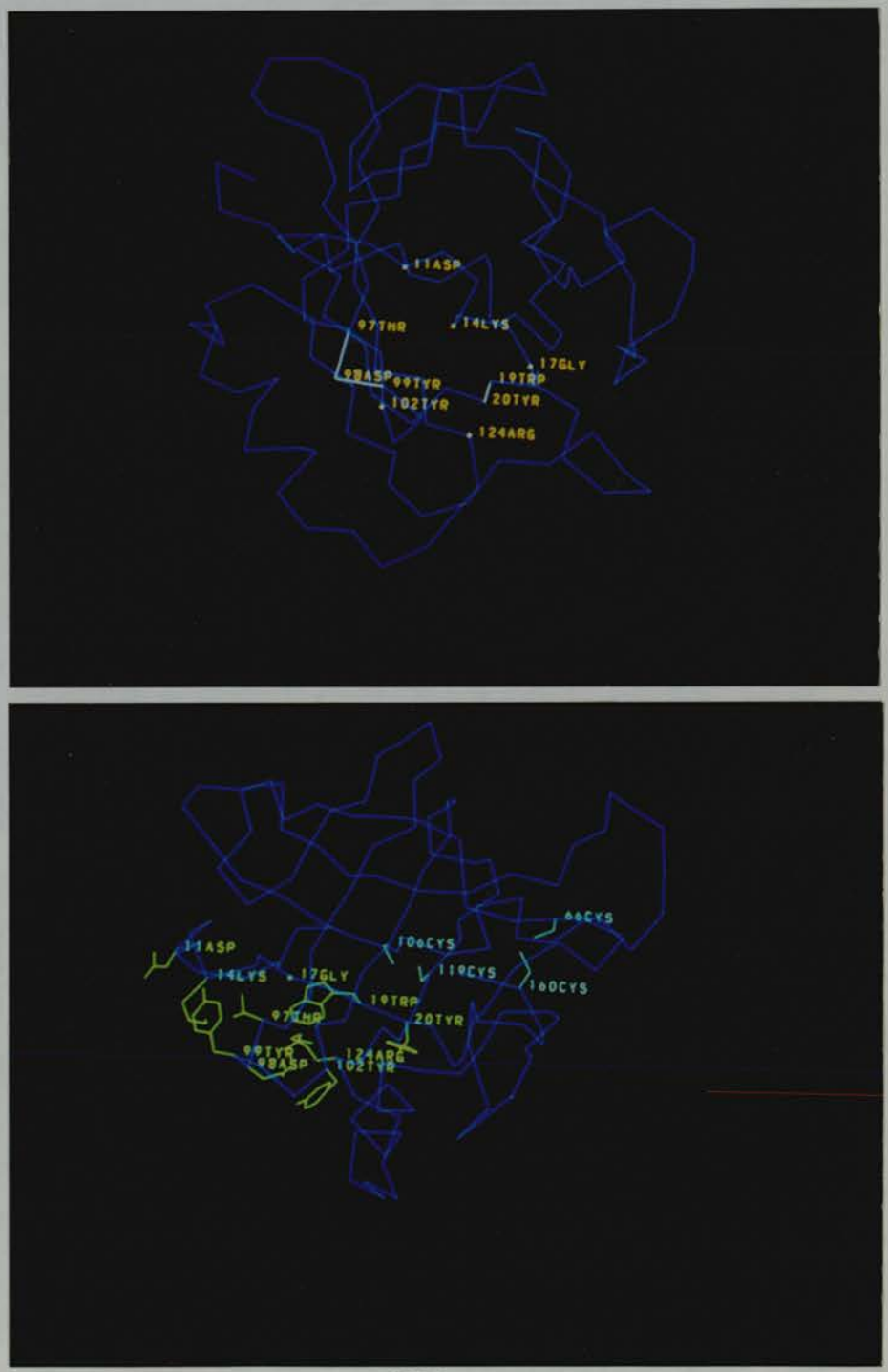


Figure 2:38

ILLUSTRATION OF THE POSITIONS OF SOME OF THE RESIDUES ASSOCIATED WITH THE β-BARREL OF BLG.

This photograph shows that Trp-19 is situated at the base of the β-barrel, Glu-51 ,Val-92 and Phe-105 are located halfway up the interior, whilst the remainder are near the mouth of the cavity.

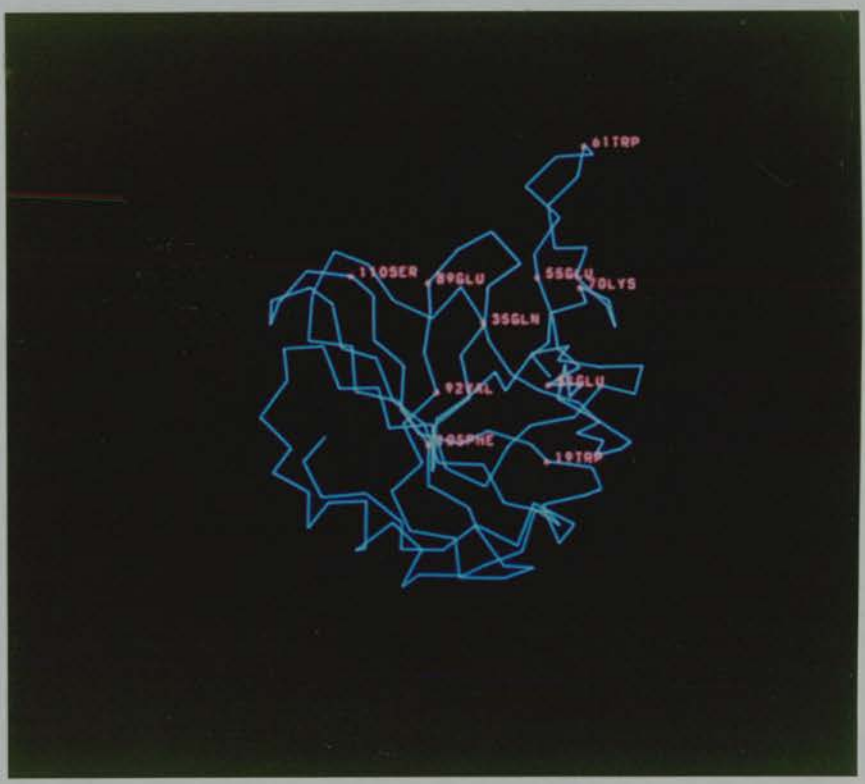


Figure 2:39

THE SITE OF ASSOCIATION OF BLG SUBUNITS UPON DIMERIZATION.

The interface between subunits in the dimeric BLG present in lattice Y involves Ile-129 and strand I (145-150). The two subunits are illustrated in yellow and blue for clarity.

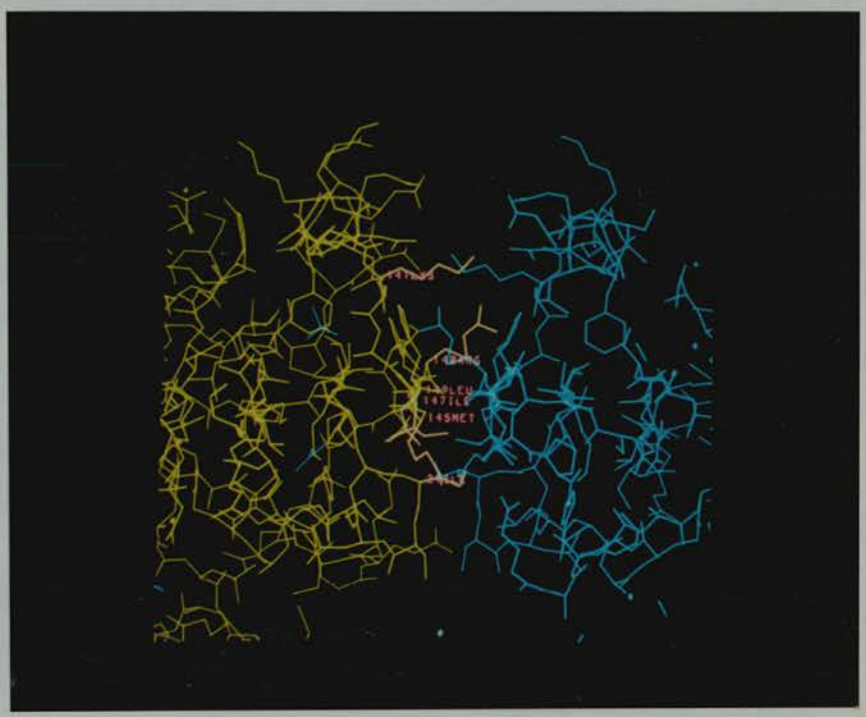
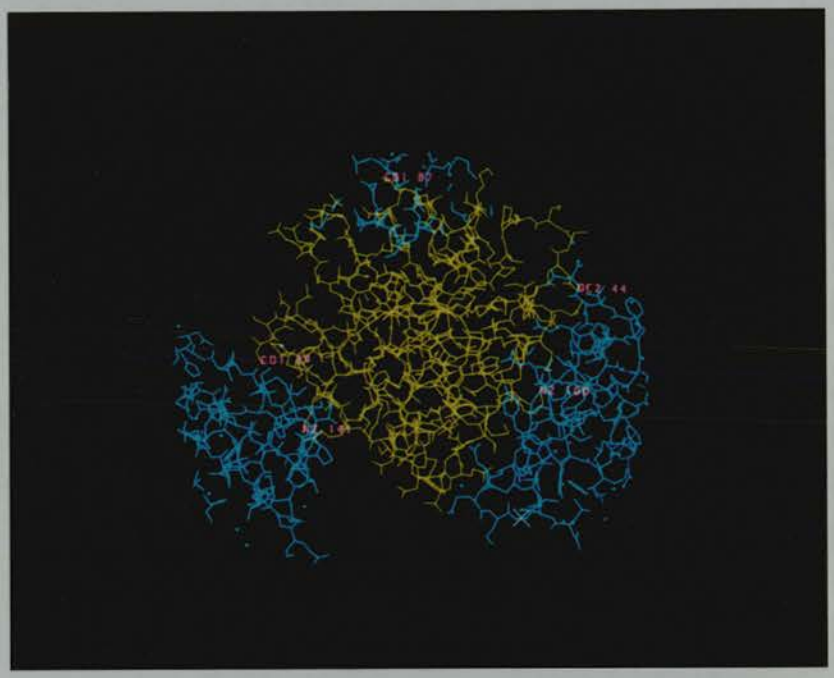


Figure 2:40

THE SITES OF ASSOCIATION OF BLG SUBUNITS WITHIN THE LATTICE Y CRYSTAL FORM.

BLG subunits not only interact hydrophobically via Ile-29 and strand I (145-150), but also electrostatically via sites around the mouth of the cavity (near Glu-89), and near residues 40-44, 100 and 124. These latter interactions are due to crystal packing effects.



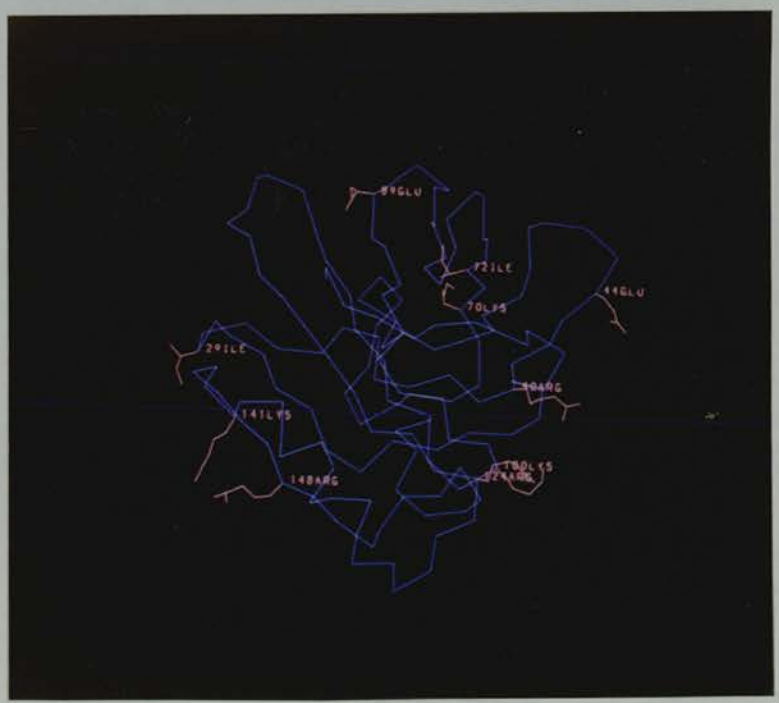
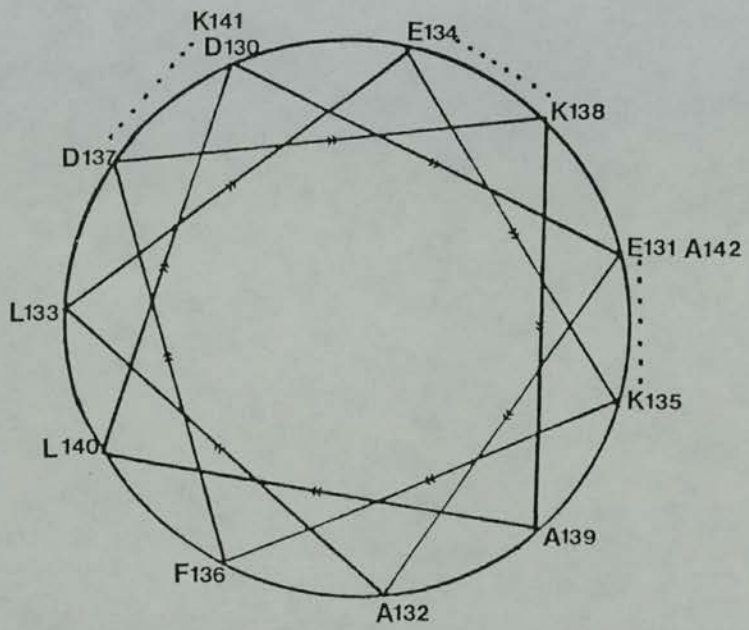
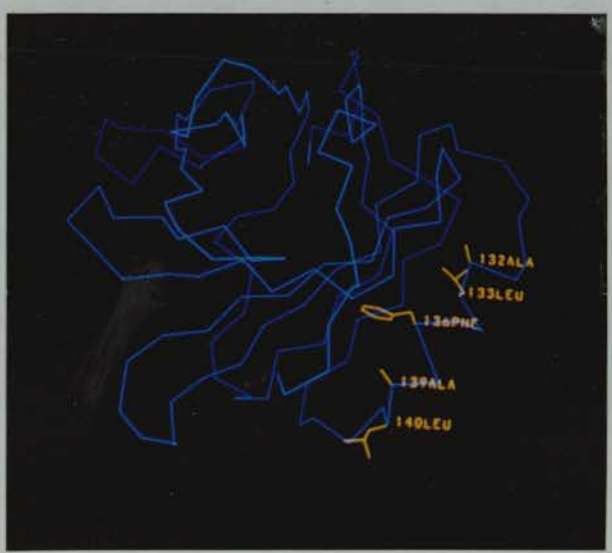


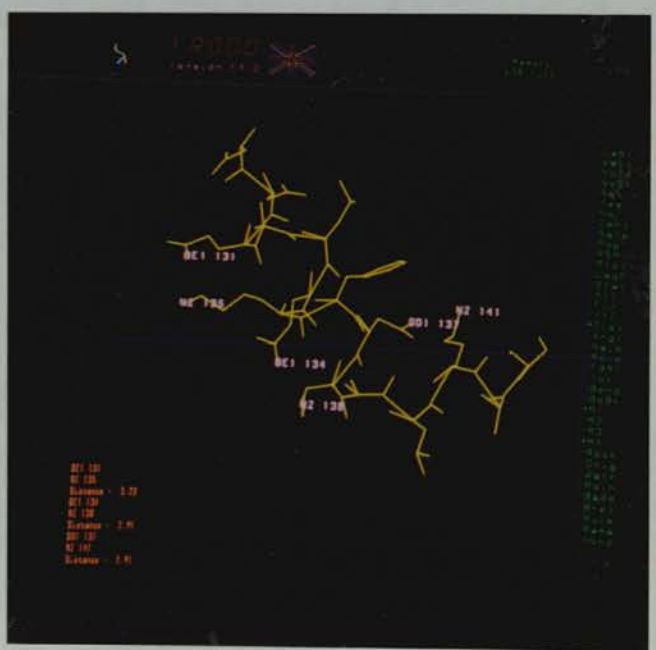
Figure 2:41

THE AMPHIPATHIC α-HELIX OF BOVINE BLG.



The below photographs illustrate that the hydrophobic residues point into the external channel, and the nature of the salt bridges formed by the hydrophilic residues.





the helix and point into the solvent, form three salt bridges: Glu-131 -> Lys-135 (3.2A), Glu-134 -> Lys-138 (2.9A) and Asp-137 -> Lys-141 (2.9A); whereas the hydrophobic amino acids Ala-132, Leu-133, Phe-136, Ala-139 and Leu-140 all point in towards the β-barrel. The space enclosed by the latter residues and strands A (16-27), H (116-124) and I (145-150), with which they interact, defines the external hydrophobic channel where retinol is bound to BLG in the lattice Z crystal [59].

2:8:3 LOOP REGIONS OF BLG LATTICE Y.

The importance of the loop regions in BLG lattice Y is difficult to determine from the structure, as their positions are often uncertain. The flexible nature of these regions indicates that it is probably their average conformation which contributes to the observed electron density, whilst their protrusion into the solvent makes them vulnerable to crystal packing effects.

The program ELLIPSE (kindly supplied by J.Thornton), was used to illustrate how far residues protruded from an ellipsoid encasing the core of the protein. For a single subunit of BLG the regions which protruded most were 28-31, 44-48, 58-67, 85-89, 110-116, 126-130, 139-147 and 156-159. However the regions 28-31 and 139-147 failed to protrude from dimeric BLG, confirming that these residues were involved in the dimer interface. All the remainder occurred in regions where turns had been fitted, or were expected.

2:8:4 THE ANTIGENIC SITES.

The antigenic sites of BLG have been studied by chemical modifications. S-carboxymethylation revealed four discontinuous sites 41-61, 62-107, 125-145 and 146-162, whereas other modifications to several specific

amino acids implicated a disulphide bridge, one Tyr, one His, some carboxyls but neither Trp nor the free thiol in these sites [209]. The probable residues involved in each site are illustrated in Figure 2:42. Three of these sites are spatially compact: 42-55, 157-161 and 127-137 - the latter is believed to be similar to the N-terminus of lactoferrin and may be where antibodies raised against lactoferrin bind to BLG [226]. The fourth site is very diverse; but as the disulphide bridge is known to be important, and there are carboxyl groups located in the locality of Cys-66, it seems likely that the residues Asp-95, Asp-96, Asp-98, Glu-74 and Glu-87 will not be involved. However it is unclear whether the regions 62-66 and 157-161 are really separate sites now that their spatial closeness has been revealed.

2:8:5 THE ENVIRONMENTS OF SOME OF THE AMINO ACIDS.

Many solution studies have previously been carried out to try and determine the environments of different amino acids. However these investigations were unable to show which specific residues were located in which environments. A detailed examination of the structure should resolve this.

(a) Trp residues.

There are two Trp residues in bovine BLG: Trp-19 and Trp-61. Trp-19 is located at the base of the hydrophobic cavity with the plane of its rings perpendicular to the axis of the β-barrel, and protected from solvent by Arg-124 which lies directly below it. This Trp is situated 28A from Trp-61 which is part of the flexible ill-defined loop 55-69, and accessible to solvent {Figure 2:43}. These observations agree well with the alkylation of BLG [155] and solvent perturbation studies [150], both of

Figure 2:42

THE ANTIGENIC REGIONS OF BOVINE BLG.

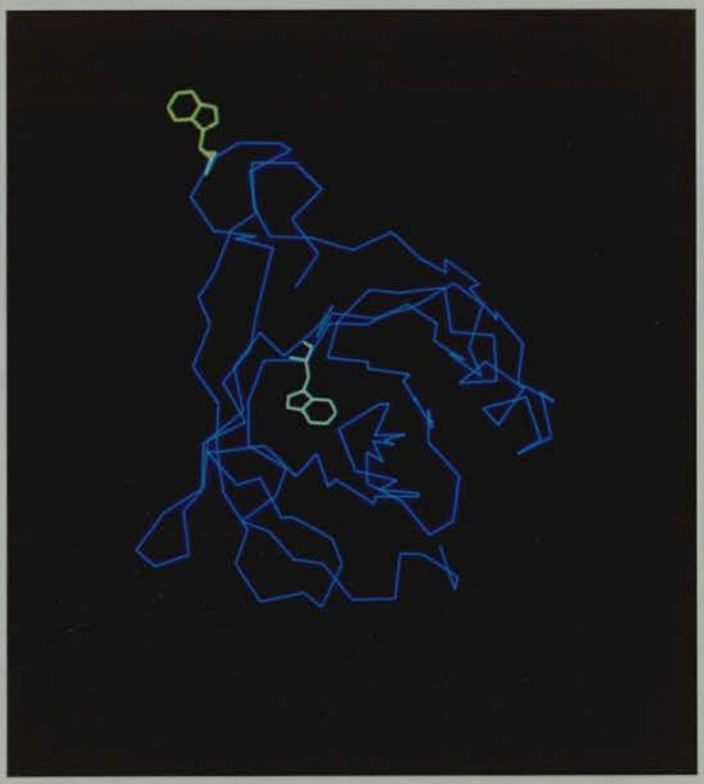
The possible residues involved in the four antigenic sites proposed from solution studies are illustrated. The four sites are in red, orange, yellow and green for clarity.

Figure 2:43

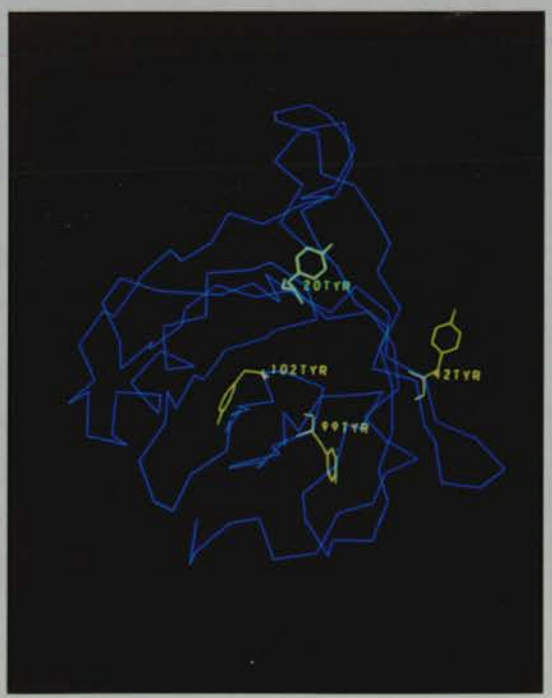
THE LOCATIONS OF THE AROMATIC RESIDUES IN BOVINE BLG.

Trp residues

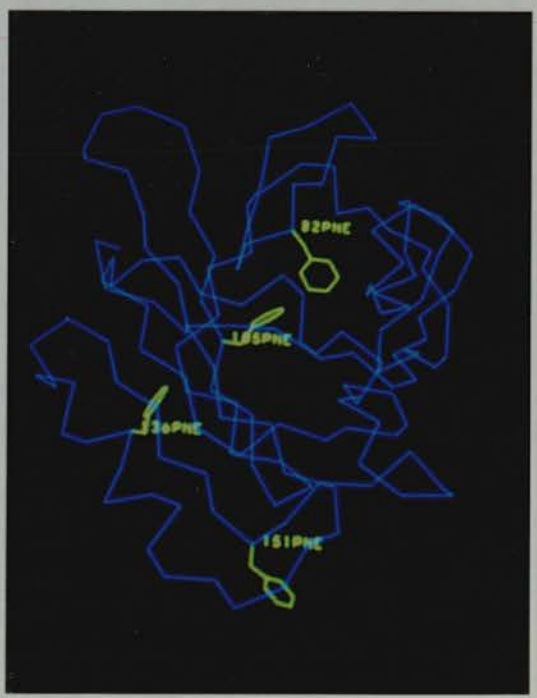
Trp-19 is located in the centre of the base of the cavity, whilst Trp-61 is situated in the extended flexible region.



Tyr residues



Phe residues



which showed that 50% of the Trp residues present were buried.

(b) Tyr residues.

Of the four Tyr residues present in bovine BLG-A or BLG-B two can be acetylated by N-acetylimidazole, with no change in the CD spectrum of the protein [150]. These surface Tyr have been identified as Tyr-20, which is located at the foot of the β-barrel on the external surface, and Tyr-42, which is near the C-end of strand B and part of one of the antigenic sites mentioned in Chapter 2:8:4 {Figure 2:43}. The reaction of cyanuric fluoride with BLG also revealed that one of the other Tyr residues was hindered whilst the fourth was buried [150]. The buried Tyr is Tyr-102 which lies enclosed between the lower region of the β-barrel and the innermost end of the surface channel. Tyr-99 is the partially accessible residue, and may H-bond to the carbonyl of either Gly-9 (4.2A) or Asp-11 (3.8A), although both distances seem rather too long at present.

(c) Phe residues.

Solution studies have not been used to determine the environments of the four Phe residues in bovine BLG. However this structure has illustrated that two of them, Phe-82 and Phe-136 are buried - the former in the hydrophobic cavity and the latter in the external hydrophobic channel. Phe-105, on strand G, appears buried half-way down within the β-barrel; whereas Phe-151 is located in the random coil at the C-terminus of the protein and is accessible {Figure 2:43}.

(d) His residues.

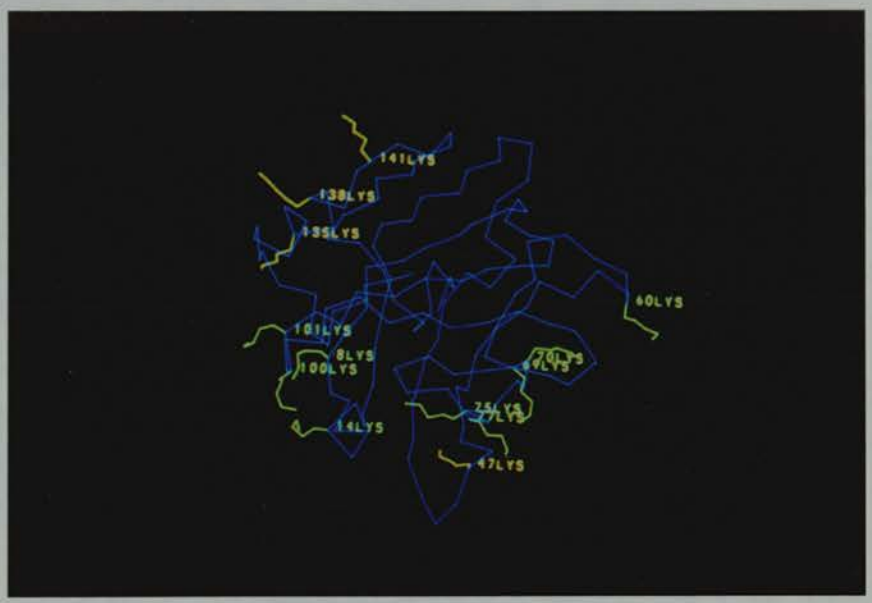
Phosphorylation of bovine BLG with $POCl_3$ at pH 8.5 enabled His, Lys and Arg residues to be modified, but could not distinguish between them [154]. Cleavege of the C-terminal residues of bovine BLG-A, His-161 and Ile-162, by carboxypeptidase-A, yielded a modified protein whose structure did not significantly differ from the native BLG [156]. This implied that His-161 was located on the protein surface, and agrees well with its location in the random-coil found at the C-terminal region. As only a single His can be modified by diethylpyrocarbonate [209] this indicates that the other His, His-146, is probably inaccessible. The structure is consistent with this - His-146 was found buried within the interface between the subunits which make up the dimeric BLG structure.

(e) Lys residues.

There are fifteen Lys residues in bovine BLG of which twelve can be covalently labelled by reductive alkylation with ^2H acetone/$NaBH_4$ at pH 9, where the protein is essentially monomeric, and then monitored by NMR. Although all the Lys residues are found on the surface of this protein structure {Figure 2:44}, four populations were detected by NMR, and can be explained in terms of the details of this structure. Lys-47 was not modified, probably due to the rigidity inflicted on it by the adjacent Pro residue; whereas Lys-135, Lys-138 and Lys-141 were only partially affected as they are part of the amphipathic α-helix, and involved in salt bridges {Figure 2:41}. Lys-69, Lys-70 and Lys-75 were all modified, but their motion was restricted by the disulphide bridge Cys-66 -> Cys-160, whilst Lys-8, Lys-14, Lys-60, Lys-77, Lys-100 and Lys-101 were also modified and free to rotate [155].

Figure 2:44

THE POSITIONS OF THE LYS RESIDUES IN BOVINE BLG LATTICE Y.



(f) Cys residues.

There are five Cys residues in bovine BLG, four of which form disulphide bridges. The bridge Cys-66 -> Cys-160 is present in all BLGs, whereas the other Cys-106 -> Cys-119 is less conserved. In pig BLG, which has no Cys-119, a disulphide bond can form between Cys-106 and Cys-121 [84]. Despite solution studies on bovine BLG which have suggested that there is an interchange between Cys-106 -> Cys-119 and Cys-106 -> Cys-121 [146,147] (but only in the presence of 8M urea which is known to affect the protein), most digestion experiments [142,143,144] revealed that Cys-121 was the free thiol. The crystallographic structures of lattices X, Y and Z all confirm this. In lattice Y the distances between these Cys residues are:

Cys-106 -> Cys 119	2.05A
Cys-106 -> Cys 121	11.56A
Cys-119 -> Cys 121	11.54A.

Thus the formation of the bridge Cys-106 -> Cys-121 appears highly unlikely, as the shift in atomic position required would necessitate the breaking of about ten H-bonds between strands G and H, and the reforming of others.

The excellent agreement between the information obtained from solution studies on the environments of specific amino acids, and the structural explanation of these details, provides strong support for the belief that the protein structure is very similar whether in solution or a crystalline form.

CHAPTER 3

THE TANFORD TRANSITION OF BLG

3:1 INTRODUCTION.

The Tanford transition is defined as the reversible change in optical rotation seen when the pH of bovine BLG varies from 6.5 to 7.8. It has been observed by both polarimetry ($[\alpha]_D$ changes from -25^o at pH 6 to -48^o at pH 8) and by ultracentrifugation ($S_{20,w}$ decreases from 3.2×10^{-13}s at pH 6 to 2.8×10^{-13}s at pH 8) {Figure 3:1} [6]. However the origin of these changes, and the relationship between this transition and the function of BLG is still not known.

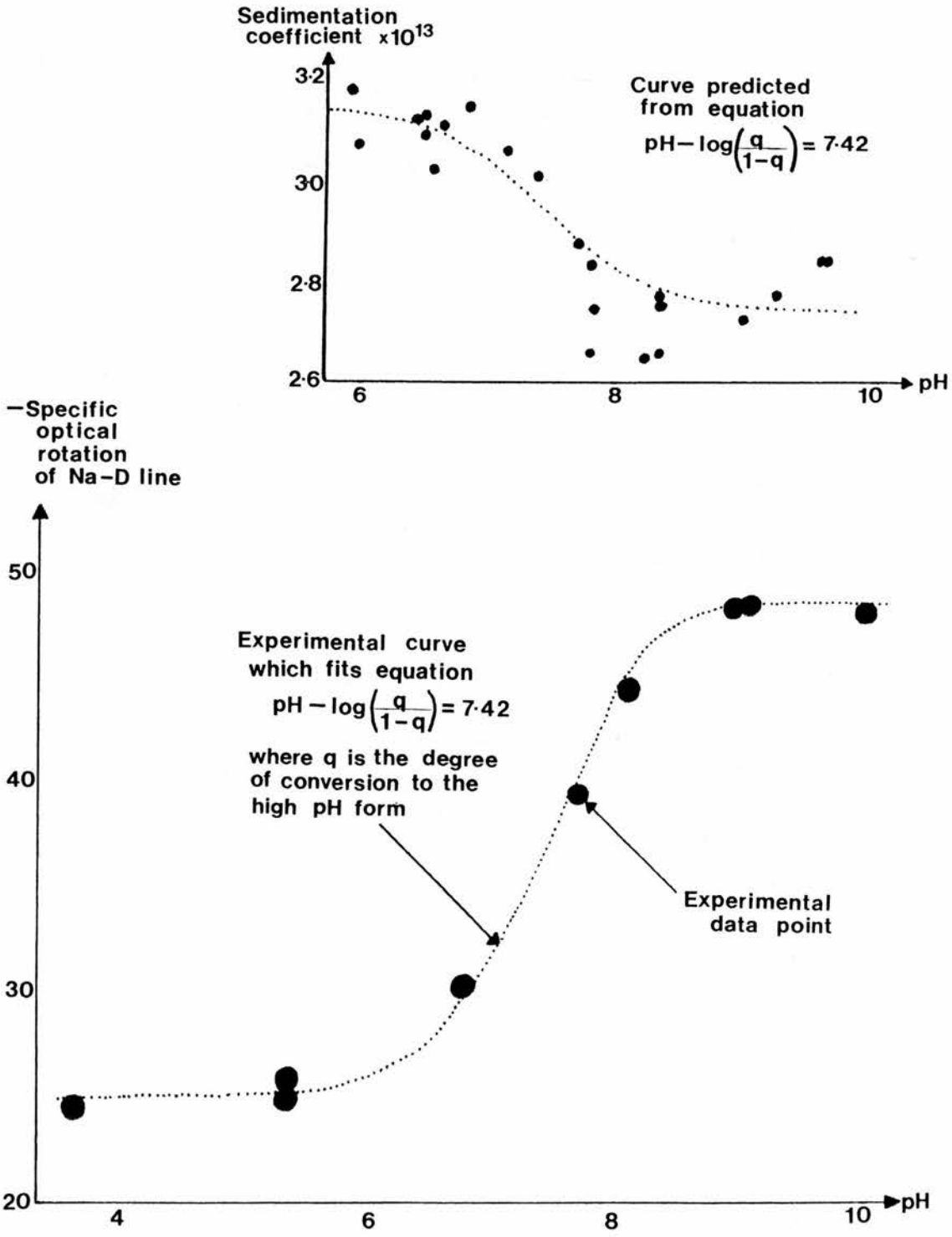
The nature of the changes observed could be due to a variety of events: a conformational change which affects the aromatic amino acids might explain the change in optical rotation, whilst the dissociation of the dimeric protein could be the reason for the decrease in sedimentation coefficient. Both possibilities are investigated in this chapter.

Far UV CD spectra were recorded at both ends of the Tanford transition to determine whether there were any gross changes in the secondary structure of the protein. Information on subtle variations in the environments of certain amino acids was obtained by near UV CD, or extracted from the literature, and related to a comparison between the crystallographic structures of lattice X at pH 6.5 and lattice Y at 7.8. It was not possible to monitor the movement of the residues thought to be involved, as the transition is known to be complete within 10 seconds [312], and stopped-flow CD was not available to us.

Although dissociation does occur over the pH range 6 to 8 it is unlikely to be a significant component of the

Figure 3:1

THE CHANGES OBSERVED IN THE OPTICAL ROTATION AND SEDIMENTATION COEFFICIENT OF BOVINE BLG OVER THE TANFORD TRANSITION.



Tanford transition for several reasons. For bovine BLG the logarithm of the equilibrium constant for monomer to dimer decreases from $5.8 l mol^{-1}$ at pH 6.5 to $4.2 l mol^{-1}$ at pH 7.5 [101], which corresponds to an increase of about 5% in the amount of monomer present. About 25% of the dimeric BLG present at pH 6.5 would need to dissociate in order to explain the decrease in $S_{20,w}$ observed. Exposure of the interface (Ile-29, Ile-147 and His-146) between subunits would also not explain the change in optical rotation seen. The increase in the volume of the protein observed upon raising the pH [38] could explain the variation in $S_{20,w}$, and seems to imply that the Tanford transition must be a conformational change. Both the variation in $S_{20,w}$ and $[\alpha]_D$ with pH have been described by the equation

$$pH - \log(q/1-q) = 7.42$$

where q is the degree of dissociation of the dimer. However this relationship appears to be coincidental, as others would also fit the rather poor sedimentation data available [6], and would be necessary to describe the larger variation in specific optical rotation (-70^o at pH 6 to -100^o at pH 8) observed when using a mercury lamp with wavelength 436nm.

The importance of the Tanford transition for the function of BLG <u>in vivo</u> is unknown. However it seems reasonable to propose that BLG may bind a ligand found in cows milk (pH 6.6), and, after ingestion by the neonate calf, transport it along the intestine. The gradual increase in pH along the GI tract may initiate the same conformational change as observed in solution, and enable the BLG complex to bind to a specific receptor. This idea is supported by several observations: the transition occurs over the physiological pH [6] and is therefore likely to be important <u>in vivo</u>, the structurally homologous protein RBP is believed to transport its ligand, retinol, in a hydrophobic cavity to a specific

receptor [7], and specific receptors for BLG have been detected in calf ileum [8]. If BLG behaves in a manner analogous to RBP [7] it should release its ligand and undergo another conformational change so as to weaken its affinity for the receptor and enable it to be replaced by another BLG+ligand complex. Whether this final conformation is related to the lattice Z structure which appears unable to bind retinol in its hydrophobic cavity is unclear.

To try and confirm the first part of this hypothesis it was necessary to prove that the Tanford transition is conserved in all BLGs, and that it is not inhibited by the binding of a ligand. Previous studies have suggested that this transition occurs in the dimeric BLGs from cow [6], goat [311], sheep [314] and red deer [38], but it has not yet been reported in monomeric BLGs. Thus polarimetric studies were carried out on monomeric pig BLG, and then on both pig and bovine BLGs in the presence of retinol, to determine whether ligand binding affected the Tanford transition.

3:2 OPTICAL ROTATION TECHNIQUES.

3:2:1 CD AND ORD TECHNIQUES.

The two techniques, circular dichroism and optical rotatory dispersion, which are related to one another by the Kronig-Kramer transform, are both based on the theory of optical rotation.

If the environment of a transition from the ground state to an excited state is asymmetric, then the displacement of charge is helical, and the molecule will be optically active. The asymmetry necessary for optical activity can originate from either the intrinsic nature of a chemical bond, or be induced by the asymmetrical

environment of a transition. In proteins intrinsic optical activity arises from the n -> π^* transition of the amide chromophore (< 240nm) and the disulphide cysteines (240-360nm), but can be modified by neighbouring groups. Induced optical activity, which is considerably weaker, about 10% of the intrinsic optical activity, may arise from the π -> π^* transition of the aromatics Trp, Tyr and Phe.

Plane-polarized light is equivalent to two circularly polarized components rotating in opposite directions. Optically active molecules interact differently with the left- and right-circularly polarized components, and this difference can be detected either as a differential change in velocity, and hence refractive index, by ORD; or as a differential change in absorption by CD {Figure 3:2}. CD spectroscopy is usually used in preference to ORD, as the bands are better resolved and more easily assigned to electronic transitions than the dispersion curves of ORD. Both techniques are restricted to only providing information on the average conformation of the molecule.

3:2:2 EXPERIMENTAL ORD.

The ORD experiments were carried out, at a single wavelength, using a Perkin-Elmer 141 polarimeter with either a mercury (λ=436nm) or sodium (λ=589nm) lamp. The former was used in all the subsequent studies as it revealed a larger change in rotation over the Tanford transition. Samples were made up by dissolving either Pentex bovine BLG or pig BLG in 0.1M phosphate buffer, and the resultant pHs were recorded on a Corning 140 pH meter.

The observed optical rotation, $[\alpha_{obs}]$, was recorded at room temperature, and then converted to the specific optical rotation via the equation

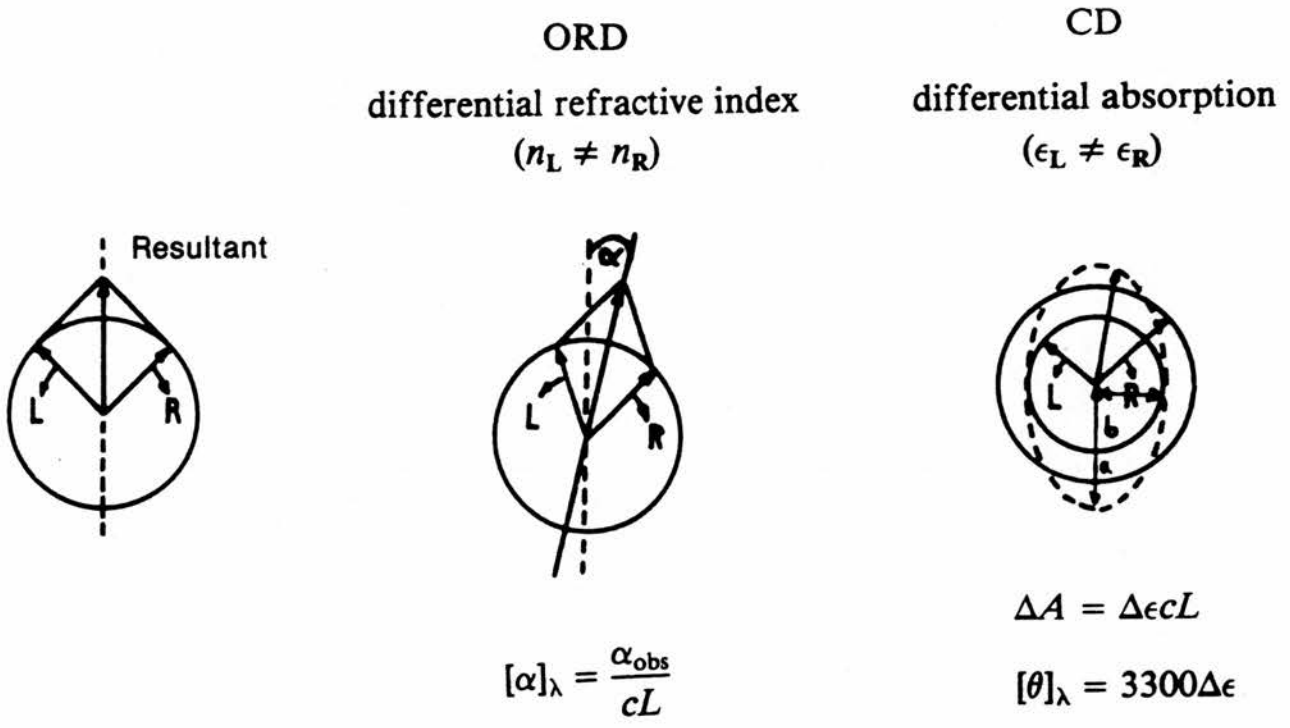


Figure 3:2

DIAGRAMMATIC REPRESENTATION OF THE THEORY OF CD AND ORD.

L and R represent the left- and right-handed components of circularly-polarized light.

$$[\alpha]\lambda = [\alpha_{obs}] / 0.999 \times c$$

where 0.999dm is the pathlength of the cell, and c is the protein concentration in g/ml. For bovine BLG $c=A_{280}/970$ where 970 is the absorption coefficient; whereas for pig BLG c was determined from dissolving a known weight of protein in a known volume. The error bars associated with the specific rotations in subsequent figures represent the standard error of the mean, calculated from three measurements.

3:2:3 EXPERIMENTAL CD.

The CD experiments were carried out at 25°C on a Jobin Yvon CD machine, which was flushed out with nitrogen prior to recording any spectra. The cells used had a pathlength ,L, of 1cm for near UV CD (260-320nm), but only 0.01cm for far UV CD (190-250nm). Air blanks were run alternately with the samples in near but not far UV CD. The samples required were usually made up on the day of use by dissolving either Pentex bovine BLG or pig BLG in 0.1M phosphate buffer, so as to obtain a A_{280} value between 0.8 and 1.2. As preliminary experiments had shown that these protein concentrations (around 1-2mg/ml) did not significantly affect the pH of the buffer, the pH values recorded were those of the phosphate buffers used. The spectra recorded from these samples were average spectra obtained from two scans.

The analysis of far UV CD spectra (190-250nm) provides information on the secondary structure of the protein. In addition to a band due to the n -> π^* transition of the amide carbonyl, there are also bands that originate from the π -> π^* transition [313]. In the case of an α-helical structure this transition becomes asymmetrical due to the presence of a large electric dipole moment being generated parallel to the axis of the helix [310]. The characteristic features for α-helix and

β-structures are:

	α-helix		β-structure	
	λ(nm)	[x']	λ(nm)	[x']
n -> π*	222	-38000	218	-8000
π -> π*	208	-35000	197	+50000
	192	+73000	190	unknown

where [x'] is the mean residue ellipticity in degree.cm^2.$dmol^{-1}$.

Near UV CD spectra (250-320nm) are useful for monitoring the environments of aromatic amino acids and disulphide bridges. The bands observed can be tentatively assigned to specific residues:

Band wavelength(nm)	Residue
254	Disulphide
260,267	Phe
275	Disulphide
275,280	Tyr/Trp
285,293	Trp

The results obtained from CD experiments can be represented either in terms of the differential absorption, ΔA, where

$$\Delta A = [x_m]cL / 3300$$

and $[x_m]$ is the molar ellipticity in degree.cm^2.$dmol^{-1}$ and is related to [x] by [x] = ($[x_m]$ / molecular weight), or by the quantity [x](mrw) where (mrw) is the mean residual weight of amino acids, equal to 112 for BLG.

3:3 MOLECULAR ASPECTS OF THE TANFORD TRANSITION.

Both the changes in optical rotation and sedimentation coefficient observed over the Tanford transition of bovine BLG [6] could be explained by a pH-induced conformational change. Various studies were carried out to determine the nature of any structural changes that occurred over the pH range 6 to 8.

3:3:1 CHANGES IN SECONDARY STRUCTURE.

The variation in the secondary structure of bovine BLG over the Tanford transition was examined by far UV CD spectroscopy. Preliminary studies were undertaken for us by P.Bayley [312], and subsequently confirmed by us {Figure 3:3}. At both pH 6 and 8 bovine BLG exhibited a negative band around 215nm and a positive signal below 200nm, which is characteristic of a β-structure, and agrees well with previously published data [175]. The differences between the two spectra were very slight indicating that there were no gross structural changes across this transition. The slight decrease in the intensity of the positive band around 195nm at pH 8 may be due to the disruption of a little of the α-helix at the higher pH.

3:3:2 AMINO ACIDS LINKED WITH THE TRANSITION.

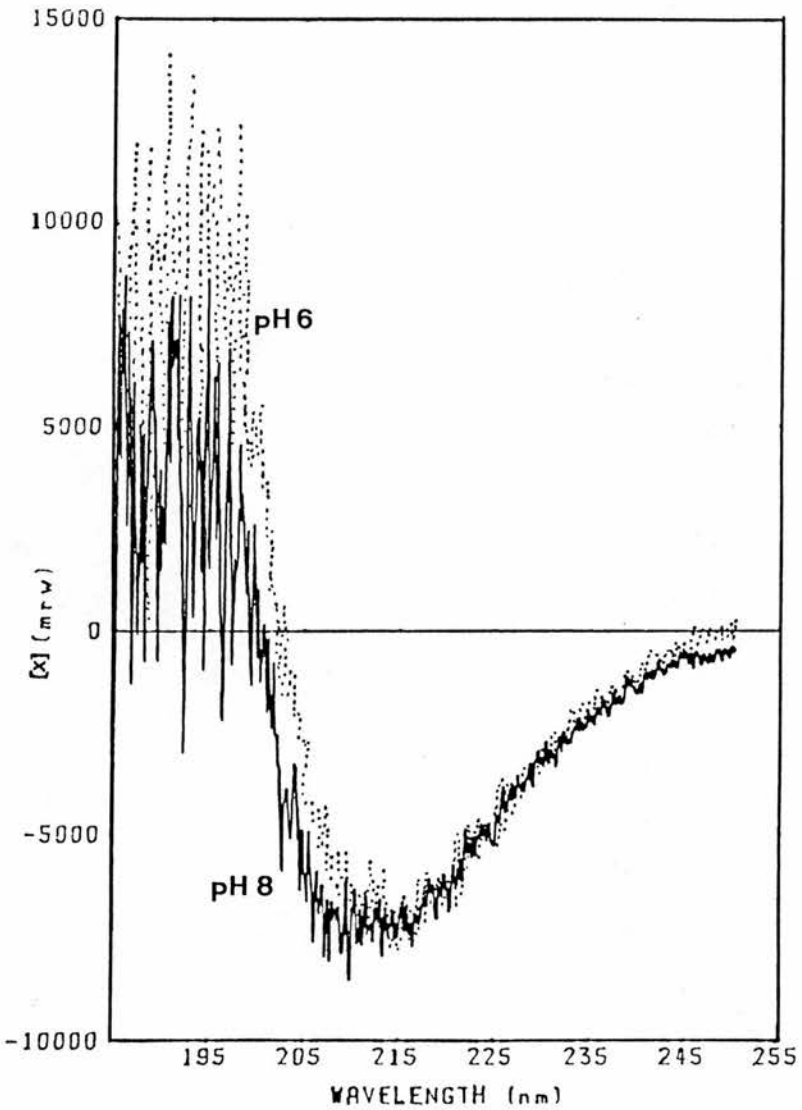
An anomalous carboxyl is involved in the Tanford transition [6], but whether its exposure and ionization is the cause or the effect is unclear. A role for His has also been suggested, but the latter is difficult to confirm. Both moieties have pK values near the midpoint of the transition (pH 7.4), and can bind Cu(II) which is known to affect the transition in bovine BLG [126].

The free sulphydryl in bovine BLG, Cys-121, is known to become more accessible to sulphydryl reagents eg. PCMB [138,139] or tetracyanoaurate(III) [140], as the pH is increased from 6.0. However it is unclear whether the free sulphydryl is involved in thiol/disulphide interchange, moves to a more exposed position, or another part of the protein undergoes the conformational change leaving it more accessible.

Difference absorption spectroscopy revealed a change in the environments of Tyr and Trp residues over the pH

Figure 3:3

FAR UV CD SPECTRA OF BOVINE BLG AT pH VALUES 6 AND 8.



range 6 to 8 [126], thus implicating them in the Tanford transition. Both residues, as well as Phe, give characteristic peaks in near UV CD spectroscopy, and hence this technique was used to try and study the role of these aromatic amino acids in this transition.

3:3:3 INVOLVEMENT OF AROMATIC RESIDUES.

Near UV CD spectra (250-320nm) were recorded on bovine BLG at pHs 6, 7.3 and 8, and these are illustrated in Figure 3:4. The residual value of [x](mrw) at 320nm was due to the presence of disulphide bands and was constant, thus helping in the superposition of the spectra. The profiles of these spectra agree well with that obtained and resolved at pH 6.4 [175]. Main bands were observed at 267, 285, and 293nm, and these have been assigned to Phe, Trp and Trp respectively, although all three contain a different, but fixed, component from the disulphide bridges [175]. Thus for each band the variation in [x](mrw) with pH must be expressed as a difference and not a ratio {Figure 3:5}. The changes measured for the bands 285nm and 293nm were considerably different, implying that they cannot both be due to Trp residues. It seems reasonable to assume that the band at 285nm is due to Tyr which normally exhibits a band around 280nm. As all three bands show a variation in their ellipticity across the Tanford transition it appears likely that Trp, Tyr and Phe residues are all involved.

3:3:4 COMPARISON OF THE BLG LATTICES X AND Y.

The types of some of the amino acids involved in the Tanford transition have been elucidated. Information on the specific residues involved could be determined from a critical look at the differences between the crystallographic structures of bovine BLG at pHs 6.5 and

Figure 3:4

THE NEAR UV CD SPECTRA OF BOVINE BLG AT THREE pH VALUES OVER THE TANFORD TRANSITION.

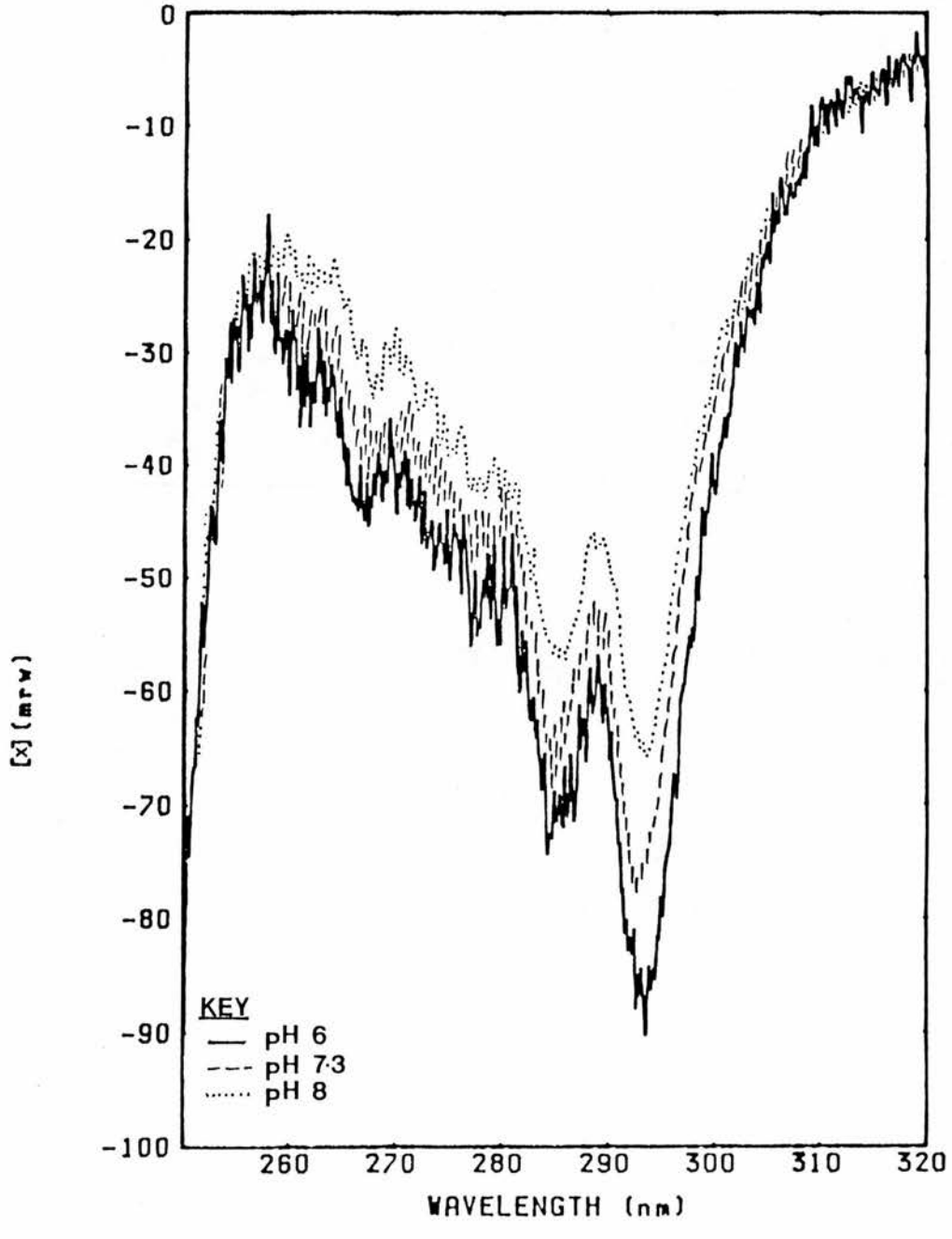


Figure 3:5

CHANGES IN THE ELLIPTICITIES, AS A FUNCTION OF pH, FOR THE MAIN BANDS IN THE NEAR UV CD SPECTRA OF BOVINE BLG.

	[x](mrw) at pH - [x](mrw) at pH 6		
	Band at 267nm (Phe)	Band at 285nm (Tyr)	Band at 293nm (Trp)
pH			
6.0	0.00	0.00	0.00
7.3	0.42	2.28	8.50
8.0	9.81	13.97	23.19

near 7.8. The pH 6.5 lattice X structure (whose determination is summarized in Appendix 5) was very kindly made available to us by S.J.Yewdall, and a preliminary comparison has been presented [283].

The superposition of lattice X on lattice Y was carried out by determining both the average rotation matrix and translation vector required to superimpose the Cα atoms of a subunit of lattice X on lattice Y. For the A subunit of lattice X the matrix

$$\begin{pmatrix} 0.35026 & 0.57984 & -0.73560 \\ -0.70802 & -0.35025 & -0.61322 \\ -0.61321 & 0.73560 & 0.28786 \end{pmatrix}$$

and translation vector (3.76074, -2.57421, 18.72389) were then applied to all the atoms in the subunit excluding residues 55-69 and 155-162. The rms distances between the A and B subunits of lattice X and lattice Y were 2.804A and 2.784A respectively, whereas that between the two subunits of lattice X was only 0.599A.

This superposition of lattice X on lattice Y showed that their secondary structures were very similar {Figure 3:6}. Slight differences were observed in the loop regions and may be due to crystal packing effects in the triclinic and orthorhombic cells. In lattice X the N-terminal loop (4-15) was more closely associated with strand F [141], whereas in lattice Y the loop at the beginning of the α-helix (125-129) was less extended, whilst that at the end (141-144) impinged less on the dimer interface.

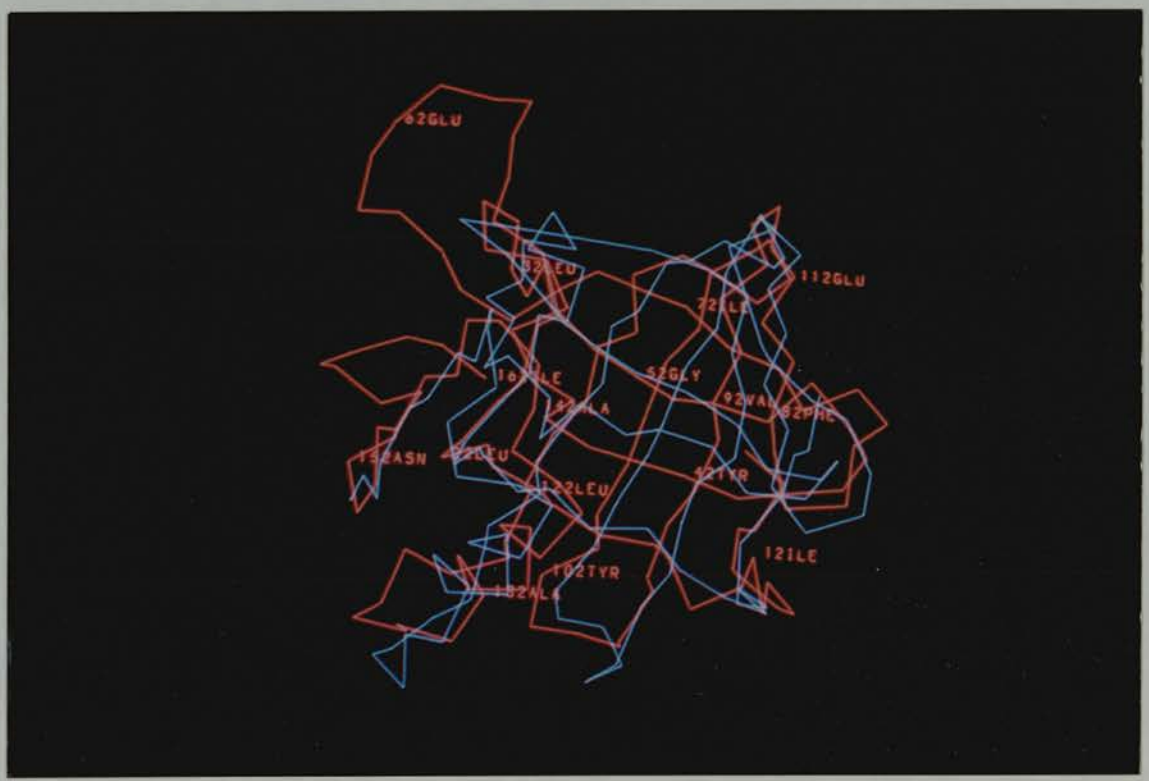
(a) Free sulphydryl.

In bovine BLG the free sulphydryl is Cys-121 in both the lattice X structure at pH 6.5 and lattice Y at pH 7.8. This implied that the reported interchange of the disulphide bridge between Cys-106 and Cys-119/Cys-121 [146,147] did not correlate with the Tanford transition.

Figure 3:6

THE SUPERPOSITION OF THE Cα ATOMS OF LATTICE X ON LATTICE Y.

Lattice Y is shown in red, whilst lattice X is in blue.



The unlikeliness of interchange was substantiated by the large separation of the sulphydryls 106 and 121: 11.6A in lattice Y and 5.9A in lattice X {Figure 3:7}. The difference between Cys-121 in lattices X and Y was 3.4A which agrees excellently with the distance between the mercurial binding site in these two structures (3.5A) [304]. The variation in the position of the free sulphydryl was due to a change in its orientation with respect to the axis of the α-helix, Cys-121 being more accessible to solvent in lattice Y. This exposure of Cys-121 across the transition fits well with the knowledge that there is about a 1000-fold increase in the reactivity of the free thiol with mercurials in the high pH structure [138].

(b) Phe residues.

Near UV CD spectroscopy has implicated the Phe residues of bovine BLG in the Tanford transition. A comparison of the two crystallographic structures revealed that all four Phe residues showed some differences between lattices X and Y {Figure 3:8}. The displacement of the residues between the lattices was small, the main difference being the orientation of the aromatic ring. The angles, through which the ring rotated between lattices X and Y, were about 50^{o}, 85^{o}, 65^{o} and 60^{o} for Phe-82, Phe-105, Phe-136 and Phe-151 respectively. The rotations of Phe-82 and Phe-105 were such that in both cases the rings protruded more into the hydrophobic cavity in lattice X than lattice Y. The most significant difference appeared to involve Phe-136, as in both structures the ring stayed almost parallel to the free sulphydryl bond {Figure 3:9}. This residue is within the α-helical region of the protein, and hence its movement over the Tanford transition could explain the slight change in secondary structure observed {Figure

Figure 3:7

THE POSITIONS OF THE FREE THIOL, CYS-121, AND THE DISULPHIDE BRIDGE CYS-106 -> CYS-119 IN LATTICES X AND Y.

Lattice X

Lattice Y

Figure 3:8

AN OVERVIEW OF SOME OF THE STRUCTURAL CHANGES OBSERVED BETWEEN LATTICES X (pH 6.5) AND Y (pH 7.8).

Lattice Y is shown in red, whilst lattice X is in green.

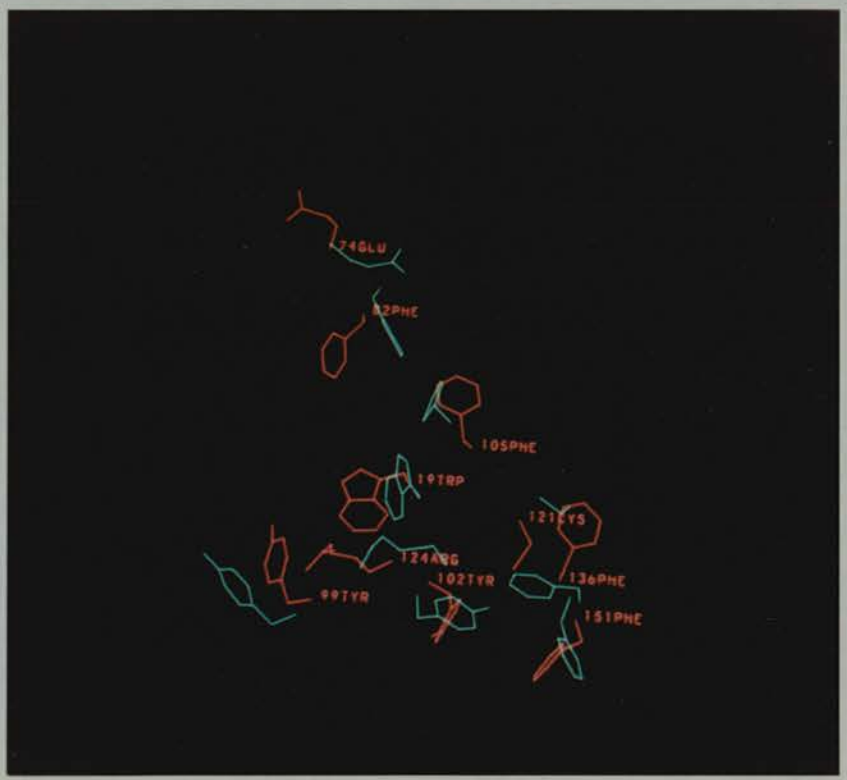
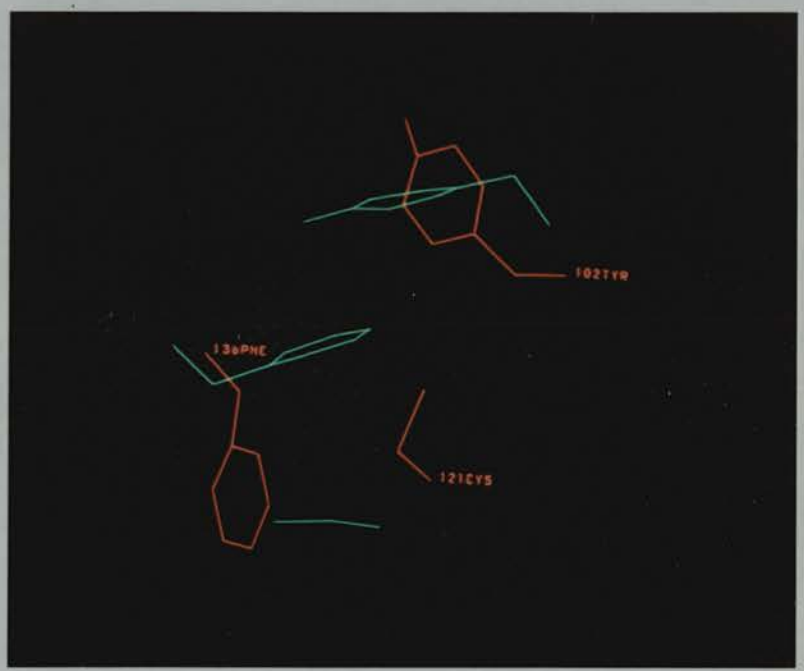
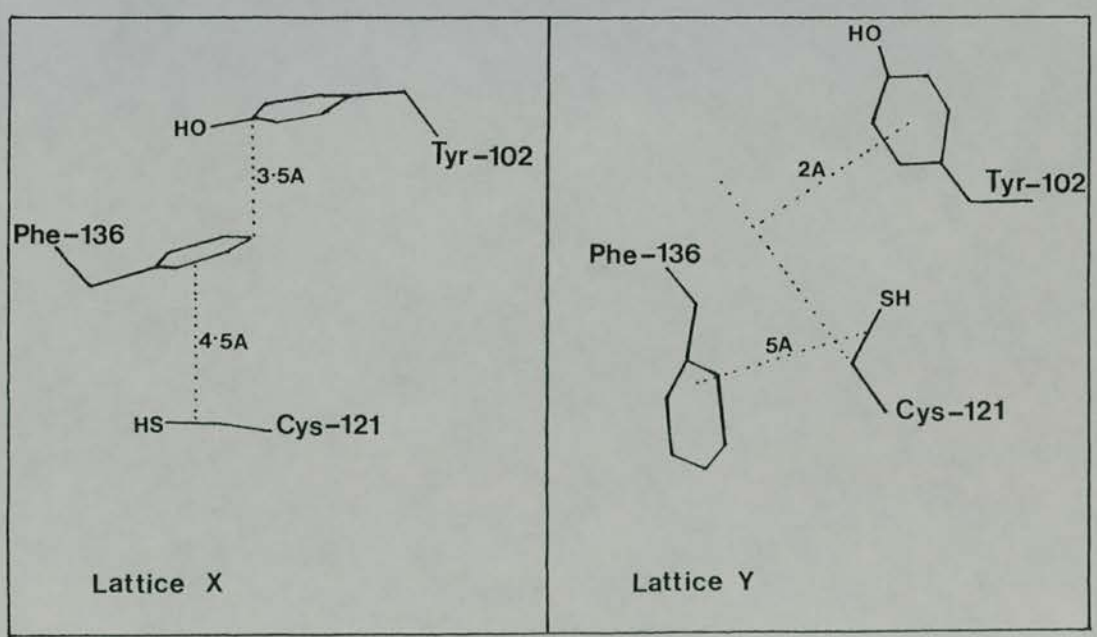


Figure 3:9

POSITIONAL DIFFERENCES BETWEEN PHE-136, TYR-102 AND CYS-121 IN LATTICES X AND Y OF BOVINE BLG.

Lattice Y is shown in red, whilst lattice X is in green.





3:3}. As the behaviour of Phe-136 over this transition seems to correlate with that of the free thiol, and it is conserved in all BLGs, it seems reasonable to propose that the Tanford transition may be linked with the function of BLG.

(c) Tyr residues.

Tyr residues had also been implicated in the Tanford transition of bovine BLG by near UV CD and difference absorption spectroscopy [126] and solvent perturbation studies. These suggested that a Tyr was unmasked at the higher pH. A comparison of the crystallographic structures of lattice X and Y have confirmed that Tyr-20 and Tyr-42 are freely accessible to solvent in both structures, their orientations differing, probably due to crystal packing effects. A slight change in the position of Tyr-99 was observed {Figure 3:8} but in both lattices it remained partially buried due to H-bonding: to the carbonyl of Gly-9 in lattice Y and of Ile-12 in lattice X. The aromatic ring of Tyr-102 changed orientation by about 85^{o} between pH 6.5 and 7.8 {Figure 3:8}. This seems likely to be significant, as Tyr-102 lies in the same spatial area of BLG as Cys-121 and Phe-136. In lattice X its ring is parallel to both the sulphydryl bond and the Phe ring, whereas in lattice Y it was no longer parallel to the sulphydryl bond, almost perpendicular to the Phe ring and more exposed to solvent {Figure 3:9}.

(d) Trp residues.

Near UV CD spectra have shown a significant change in the environment of Trp residues over the Tanford transition {Figure 3:4}. As Trp-61 is located in an ill-defined flexible loop its position is poorly known in both lattices X and Y. Thus a knowledge of its role in

this transition, if any, must await improvements in the crystallographic structures. Trp-19 is located at the foot of the hydrophobic binding cavity in both structures. In lattice Y the indole ring is perpendicular to the axis of the barrel, and protected from solvent by the side chain of Arg-124 which lies directly below it; whereas in lattice X the ring is inclined at about 20° to the barrel axis and Arg-124 is displaced towards strand A. The increased exposure of Trp-19 to solvent at pH 6.5 relative to 7.8 may be linked with the Tanford transition. As Trp-19 is also believed to be involved in the binding of retinol to BLG [153,160], this suggests that there could be a link between this conformational change and the binding of ligands to BLG.

(e) Anomalous carboxyl.

The single anomalous carboxyl which is exposed and ionized per subunit of bovine BLG, over the Tanford transition, is more difficult to identify. Bovine BLG-A has 27 potential candidates: 11 Asp and 16 Glu residues, some of which can be excluded.

Glu-131, Glu-134 and Asp-137, which are part of the amphipathic α-helix, are exposed to solvent and involved in salt bridges in both lattices X and Y, thus eliminating them from being the anomalous carboxyl. Residues located in the middle of the flexible loops, or the N-terminal and C-terminal regions, are unlikely to be involved as their motions are less restricted, and hence they will tend to be exposed in both lattices eg. Asp-11, Glu-112, Glu-114, Glu-157 and Glu-158. A comparison of the primary sequences of bovine BLGs A, B, C, D and goat BLG, which are all known to exhibit the Tanford transition [104,74,311], revealed some residues which could be eliminated. Glu-45 was not present in BLG-D and thus not the anomalous carboxyl; whereas Asp-64 was

absent in BLG-B but could not be excluded, as this variant contains the sequence Asn-63 Gly-64, the latter being able to deaminate the former effectively giving 'Asp-63' [319]. The slight differences between the environments of Asp-64 in BLG-A and 'Asp-63' in BLG-B could explain the small variations observed in the optical rotations of these proteins [104]. This cannot be confirmed at present, as the positions of these residues are poorly defined in both lattices X and Y. Goat BLG does not possess Asp-53, Asp-130 nor Glu-158, unlike bovine BLGs, thus implying that these residues are not the cause of the anomalous behaviour. Asp-53 was firmly ruled out from a comparison of the X and Y structures - it lies on the inner surface of the calyx and is partially shielded from solvent by Gln-35 in lattice Y, but is not significantly more buried in lattice X [141].

A comparison of the lattice X and lattice Y structures also revealed three carboxyls, Asp-33, Glu-74 and Asp-96, which were located in the interior of the protein and involved in H-bonding at pH 6.5, but were in different, more exposed positions around pH 7.8. The positions of these residues in both structures, and an indication of the H-bonds which they could form in lattice X, but not in lattice Y, are given in Figures 3:10 and 3:11.

There is evidence to support each candidate. Glu-74 has been proposed as the anomalous carboxyl previously [141], but no evidence was offered as to why the other two residues were unsuitable. It is conserved in bovine BLG-A, -B, -C, -D, and goat BLG, but is substituted by Lys in pig BLG. This indicates that if Glu-74 is the anomalous carboxyl then the conformational change cannot be initiated by the exposure of the carboxyl. Asp-33 could also be the anomalous carboxyl from the variation in its postion between lattices X and Y. Asp-96 seems to be the most feasible candidate to show anomalous

Figure 3:10

THE POSITIONS OF ASP-33, GLU-74 AND ASP-96 IN LATTICES X AND Y OF BOVINE BLG.

Lattice Y is shown in red, whilst lattice X is in blue.

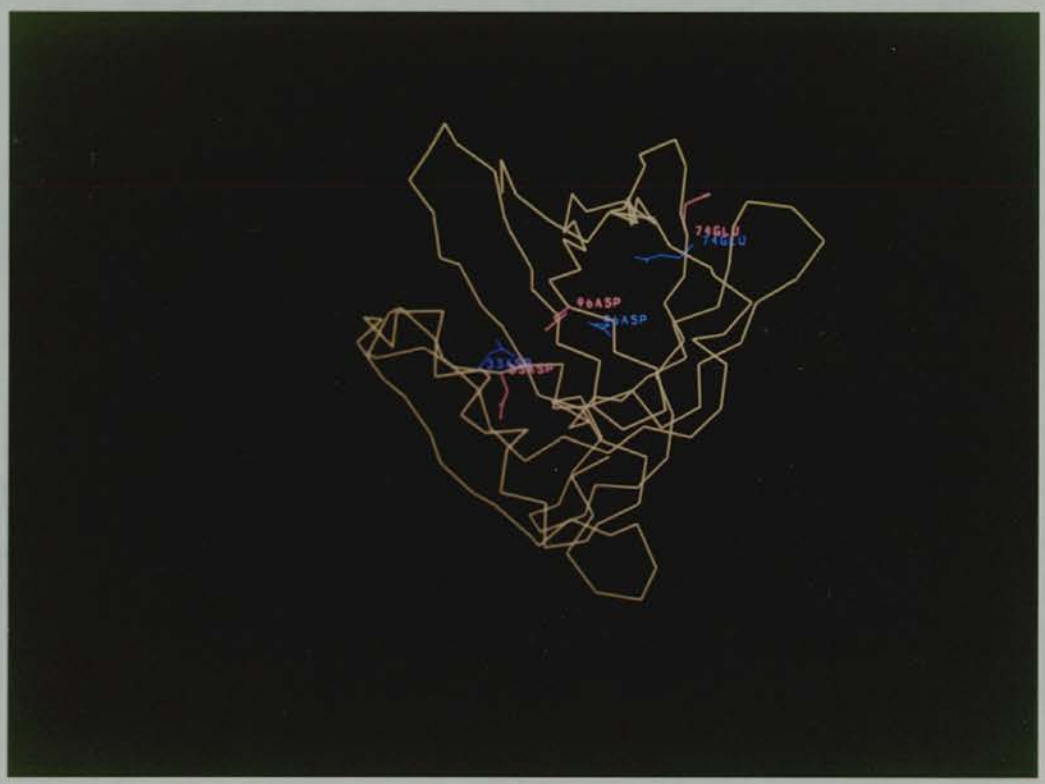


Figure 3:11

POSSIBLE H-BONDING INTERACTIONS INVOLVING ASP-33, GLU-74 AND ASP-96 IN LATTICE X, THAT ARE ABSENT IN LATTICE Y.

Interaction	Distances in A	
	Lattice X	Lattice Y
OD1-33 -> OE1-35	3.47	9.25
OD2-33 -> O-31	3.44	7.53
OE1-74 -> OE1-51	4.14	13.23
OE2-74 -> O-82	2.56	8.27
OD1-96 -> O-102	2.85	9.09
OD2-96 -> O-95	3.02	4.48

Symbols.

OD1-n : The first delta oxygen atom in residue n.
OD2-n : The second delta oxygen atom in residue n.
OE1-n : The first epsilon oxygen atom in residue n.
OE2-n : The second epsilon oxygen atom in residue n.
O-n : The carbonyl oxygen atom in residue n.

behaviour, as it situated in the same region of the protein as Phe-136 and Tyr-102, and is conserved in all BLGs except kangaroo which has Glu-96 (which could also behave anomalously) and horse BLG-II that has Pro-96. The latter protein seems an 'odd' BLG as it shows a very low sequence homology to bovine, caprine or porcine BLGs and does not contain the regions conserved in all other BLGs. As site-directed mutagenesis of the ovine BLG gene can now be carried out the most promising way to unambiguously solve the identity of the anomalous carboxyl seems to be the replacement of each of these candidates in turn, and then polarimetric studies of the mutated proteins.

(f) His residues.

Although a role for His residues in the Tanford transition has been proposed (they have a pK value of 7.3, almost equivalent to the midpoint of the transition, and can bind Cu(II) which affects this transition) this seems unreasonable. His-161 is not likely to be important, as carboxypeptidase-A cleavage of the penultimate residues of bovine BLG does not affect the CD spectrum of the protein [156]; whilst His-146 is unlikely to be involved as it is buried in the dimer interface of both lattices X and Y. The comparison of lattices X and Y revealed insignificant changes in the position of His-146, whilst His-161 could not be investigated as its location in lattice X is unknown.

3:4 EVIDENCE FOR THE TANFORD TRANSITION IN VIVO.

3:4:1 OCCURRENCE OF THE TANFORD TRANSITION IN PIG BLG.

If the Tanford transition is important for the

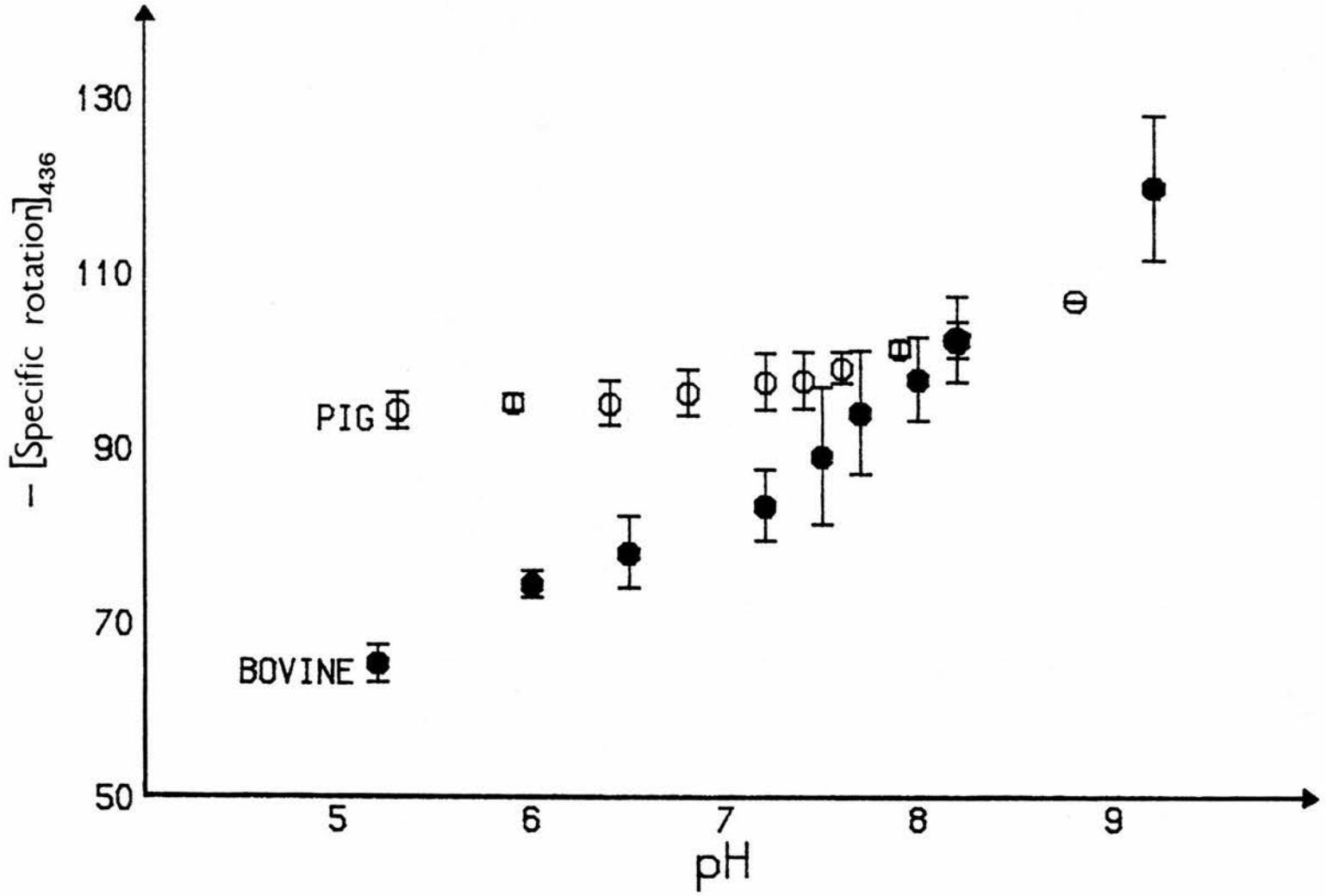
biological function of BLG, it should be present in all BLGs. ORD studies have revealed that this transition occurs in bovine [6] and goat BLGs [311] - this was not surprising as they are both dimeric proteins with very similar primary structures {Figure 1:10}. However it has also been suggested that monomeric pig BLG shows a change in optical rotation over the physiological pH, of about a third of that observed for bovine BLG [314].

To check this unpublished result the optical rotation of pig BLG over the pH range 6 to 8 was recorded by polarimetry {Figure 3:12}. The difference in $[\alpha]_{436}$ across this range was 7^{o} for pig BLG, as compared to 23^{o} for bovine BLG, thus confirming the above observation. The cause of this smaller variation in $[\alpha]_{436}$ could be due to the lack of a contribution from the dissociation of the protein (pig BLG is monomeric over the pHs studied), a difference in the secondary structure between pig and bovine BLGs, or the substitution, in pig BLG, of amino acids that have been implicated in the Tanford transition of bovine BLG.

To investigate whether the dissociation of bovine BLG contributed to the change in optical rotation observed over this transition it was necessary to dissociate the two subunits of bovine BLG whilst maintaining their native structure. The literature reported that the addition of more than 3M urea to a BLG solution would cause the formation of monomers, whilst the addition of up to 4M urea would not affect the native structure, as detected by ORD [119]. Thus Pentex bovine BLG was dissolved in a solution of 4M urea in 0.1M phosphate buffer at pH 7.3. The effect of 4M urea on the near UV CD spectrum of bovine BLG at pH 7.3 is given on the next page.

Figure 3:12

THE TANFORD TRANSITION OF PIG AND BOVINE BLGS.



Wavelength of band (nm)	[x](mrw) Bovine BLG	Bovine BLG+4M urea
320 (Disulphide)	-5	-5
267 (Phe)	-43	-28
285 (Tyr)	-70	-46
293 (Trp)	-79	-44

The constant ellipticity at 320nm confirmed that the disulphide bridges had not been disrupted by urea, whereas the changes observed at 267, 285, and 293nm revealed that there were changes in the environment of the native structure. Hence this method was not suitable for obtaining subunits of bovine BLG with a native conformation. Recently site-directed mutagenesis has produced a mutated ovine BLG with Ser-29 instead of Ile-29 [93]. This protein is known to be monomeric, but its ability to exhibit the Tanford transition has not yet been investigated.

The secondary structure of pig BLG was examined at both pH 6 and 8 by far UV CD spectroscopy {Figure 3:13}. These spectra appeared very similar: no signal was measured below 200nm due to noise, whilst a negative band was observed at 207nm with a minor shoulder at 222nm. The shoulder was slightly enhanced at pH 6, indicating that there was a small increase in the α-helical content of the protein at the lower pH. This behaviour is analogous to that of bovine BLG. However a comparison of the pig and bovine BLG spectra, at either pH, revealed that their structures were different {Figures 3:3 and 3:13}. The displacement of the negative band from 215nm in bovine to 207nm in pig, together with the exposure of a weak shoulder at 222nm in the latter protein, implies that there is a larger amount of α-helix in pig BLG.

The near UV CD spectra of pig BLG were recorded at pHs 6, 6.5, 7.3, 7.8 and 8. Three of these are illustrated {Figure 3:14}, the other two, at pH 6.5 and

Figure 3:13

FAR UV CD SPECTRA OF PIG BLG AT pH VALUES 6 AND 8.

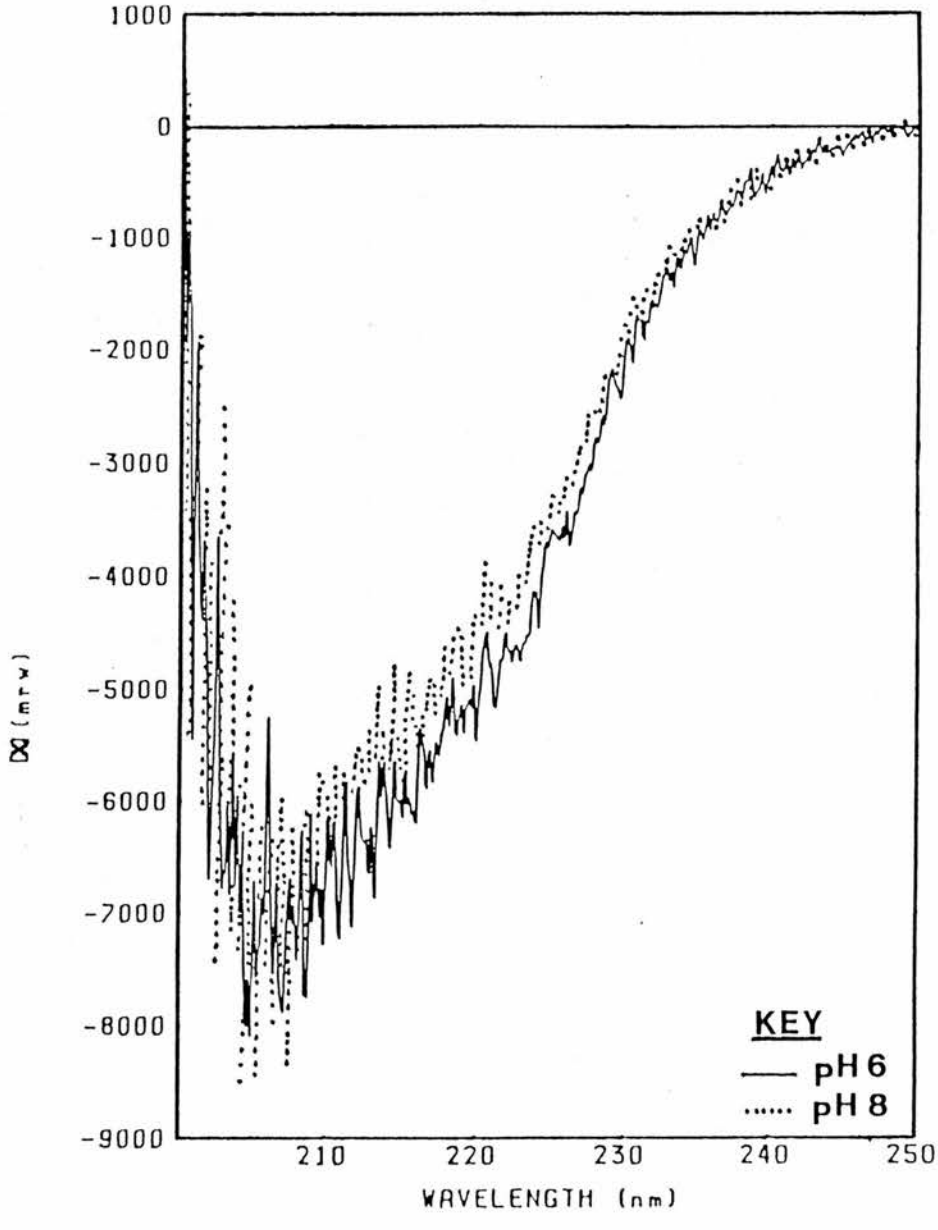
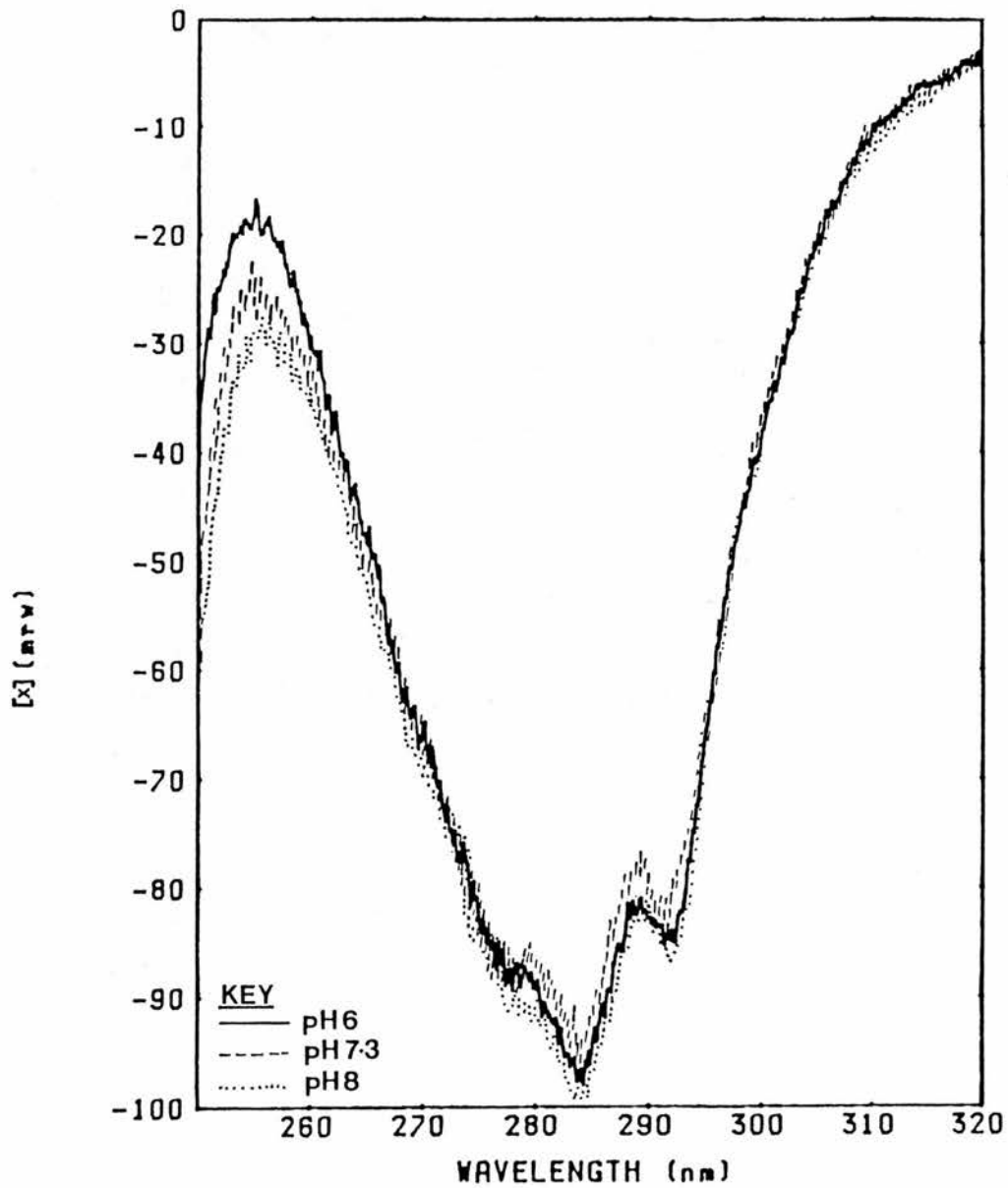


Figure 3:14

THE NEAR UV CD SPECTRA OF PIG BLG AT THREE pH VALUES OVER THE TANFORD TRANSITION.



7.8, were both exactly superimposable upon the curve for pH 7.3. The profile of the spectra differed considerably from those of bovine BLG {Figures 3:4 and 3:14}, which was not surprising as these two proteins had slightly different structures, and exhibited only 61% amino acid homology {Figure 1:11}, and had different numbers of Tyr and Trp residues. The residual [x](mrw) value at 320nm remained constant, around -5, suggesting that pig BLG also contains two disulphide bridges, Cys-66 -> Cys-160 as in bovine, and Cys-106 -> Cys-121 because Cys-119 is absent in pig BLG.

The substitution of Tyr-20 and Tyr-102 in pig BLG may explain why the spectra do not vary much over the pH range 6 to 8. The absence of Tyr-20 may explain the constancy of the 293nm band due to its neighbour Trp-19; whilst the lack of Tyr-102, which has been implicated in the Tanford transition of bovine BLG, may explain why the band at 285nm remains constant. Its absence, together with that of any free thiol, may cause a change in the environment of Phe-136, and produce the slight variation in the signal around 257nm. However, until crystallographic structures of pig BLG are available at each end of its modified transition, any structural observations must remain tentative.

3:4:2 EFFECT OF BINDING ON THE TRANSITION OF BLGS.

If, as proposed, the Tanford transition is a pH-induced conformational change which BLG undergoes in the GI tract prior to its binding to a specific receptor in the lower ileum, then it must not be inhibited by the binding of its ligand. Retinol has been proposed as the true ligand for BLG because it is found in milk at a concentration of about 1mg/ml, forms a 1:1 complex with bovine BLG that has a K_d of 0.02μM [153], and can bind to the structurally homologous protein RBP [215].

Polarimetric studies were undertaken to determine the effect of retinol on the Tanford transition of both pig and bovine BLGs. The optical rotation of each protein sample was measured before and after a small amount of a concentrated retinol/EtOH solution was added, such that the concentration of BLG remained constant and there was a slight excess of retinol. The percentage of EtOH present was kept to only 1%, as this level had been shown not to alter the optical rotation of the protein, and its effect on the pH of the sample was known and corrected for.

The data for both bovine and pig BLGs are illustrated in Figures 3:15 and 3:16. The addition of retinol to bovine BLG does not affect the profile of the Tanford transition, but does appear to displace the transition towards lower pH values, the midpoint decreasing from pH 7.5 to 7.2. It is interesting to note that the latter pH is nearer that at which there is maximal uptake of retinol from the BLG+retinol complex by specific receptors <u>in vitro</u> {Figure 5:2}. However the addition of retinol to pig BLG does not seem to affect the transition. There are several possible explanations. The retinol may not bind to pig BLG, which would imply that it is not the true ligand for BLGs, and this will be investigated in Chapter 4. Alternatively pig BLG may bind retinol either at a different site compared to bovine BLG, or a similar one providing that one or more of the residues present are those implicated in the Tanford transition of bovine BLG, but substituted in pig BLG. As these substitutions reduce the transition by 70%, they may also reduce the effect of ligand binding on the transition - a displacement of 30% of 0.3 pH units would be too small to be clearly detected. The nature of the retinol binding to both bovine and pig BLGs was therefore examined in further detail in Chapter 4.

Figure 3:15

THE EFFECT OF RETINOL ON THE TANFORD TRANSITION OF BOVINE BLG.

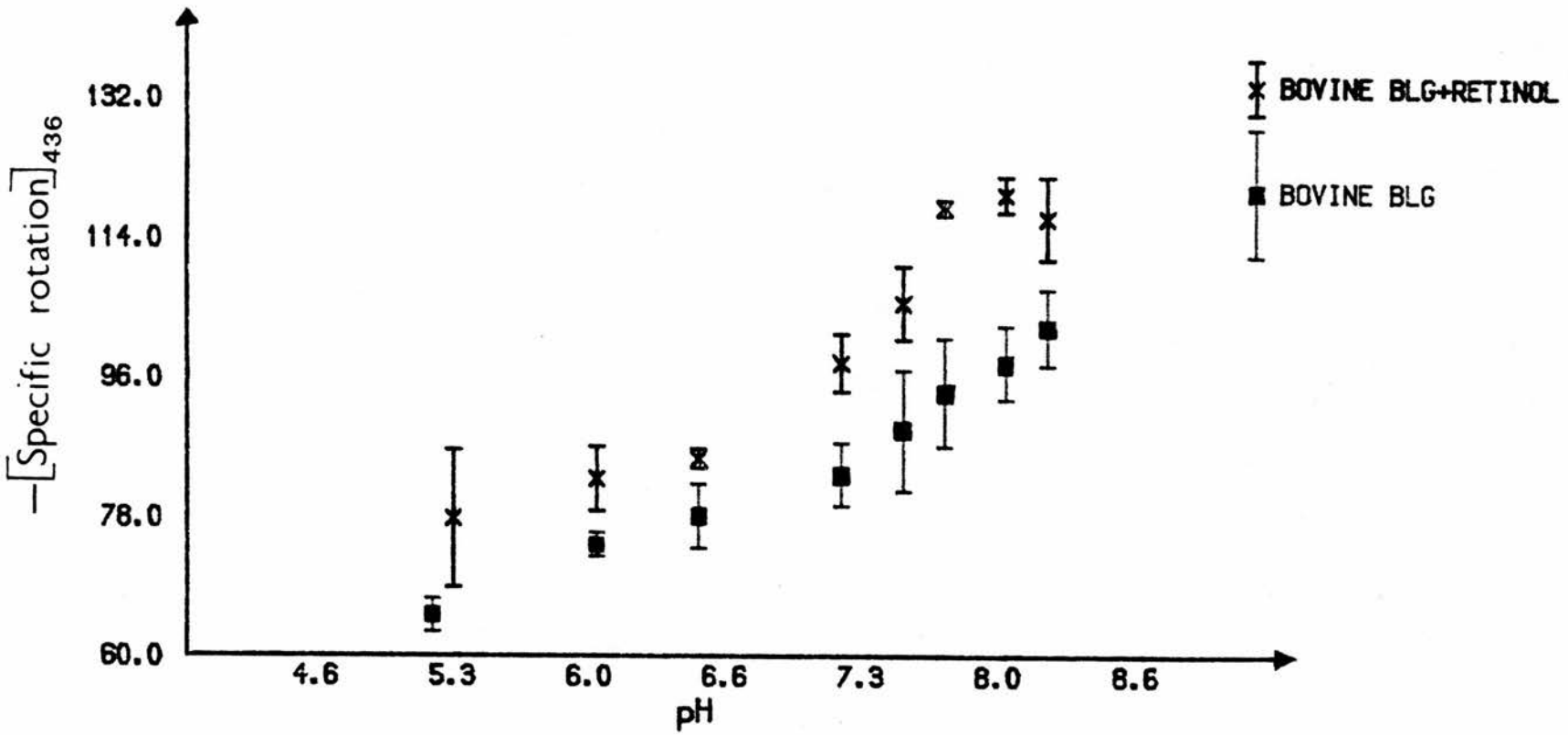
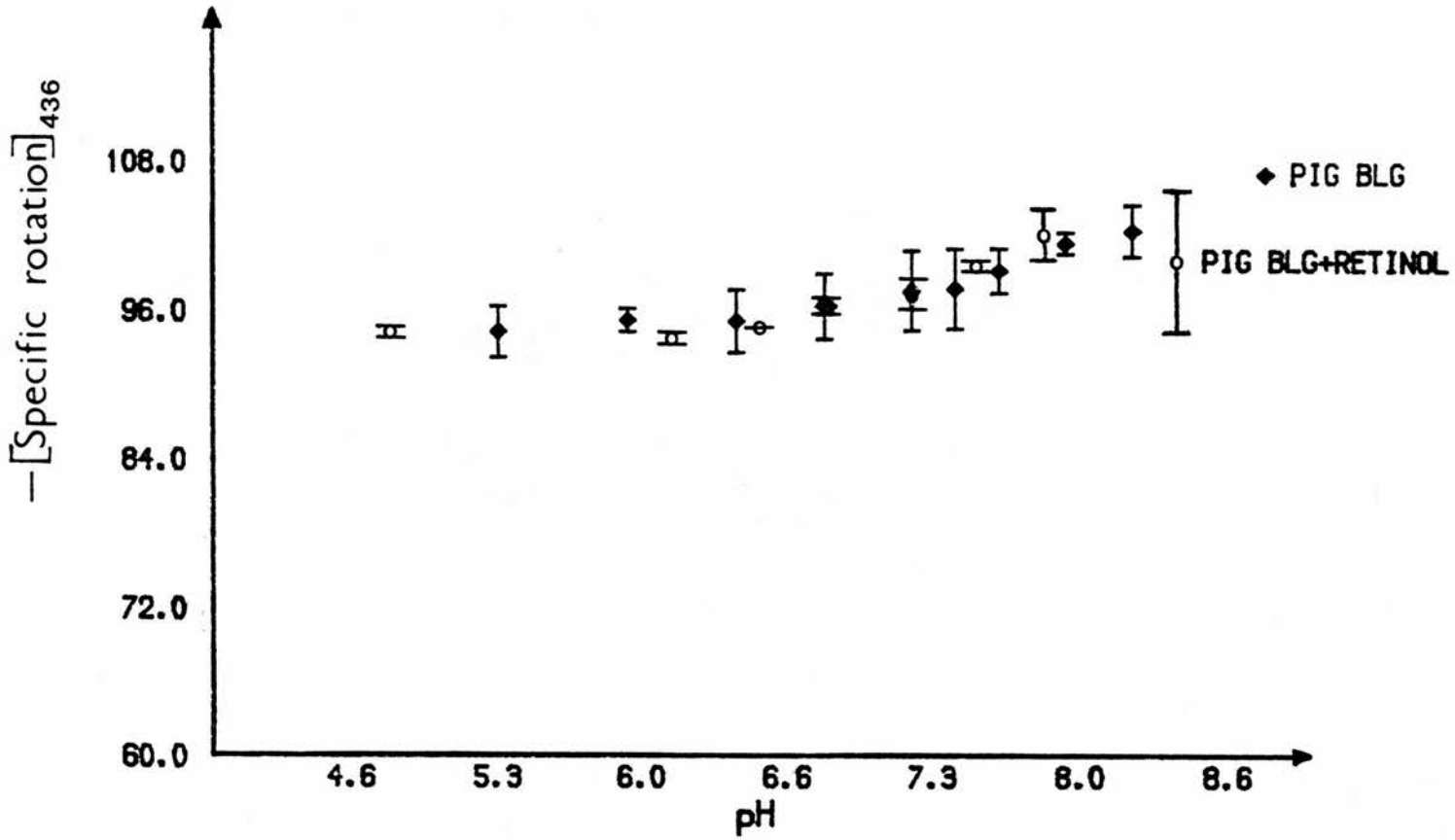


Figure 3:16

THE EFFECT OF RETINOL ON THE TANFORD TRANSITION OF PIG BLG.



CHAPTER 4

LIGAND BINDING TO BLG

4:1 INTRODUCTION.

An assortment of small molecules is known to bind to bovine BLG in solution {Chapter 1:7}. As the nature of these ligands varies considerably, from aromatic hydrophilic compounds like PNPP [164] to hydrophobic substances including retinol [153,160], it seems likely that this protein may have more than one binding site, and that these may be non-specific. The binding of 2,6-MANS to bovine BLG revealed two binding sites, a single strong site with an association constant of $3.4 \times 10^5 M^{-1}$ and a weaker one. Both were present at each end of the Tanford transition [162]. However details on the nature of these sites are sparse.

One binding site is thought to be in the external hydrophobic channel between the free thiol, Cys-121, and the α-helix. Retinol apparently binds here in the lattice Z crystal of bovine BLG, whilst IEBA has been associated with the equivalent region in RBP [217]. The importance of this site is unclear, as Cys-121 is not conserved in all BLGs, but is implicated in the Tanford transition. As the binding of ligands at this site may affect the transition polarimetric studies were undertaken.

The other site is believed to be within the hydrophobic cavity of the β-barrel of BLG, being proposed from the location of retinol in the structurally homologous protein RBP [215]. Trp-19 has been implicated in this binding site [8]; and, as it is conserved in all BLGs, and within the superfamily of homologous proteins {Figures 1:10, 1:11 and 1:15}, is likely to be important for the function of BLG, and have a role in

transportation. The effect of various ligands on this binding site was investigated by Trp fluorescence and near UV CD spectroscopy.

In the subsequent ligand binding studies it was assumed that both binding sites were free from any endogenous ligand. At the outset of these experiments this seemed reasonable, as the proteins had been isolated by procedures involving high ionic concentrations and large variations in pH, and no unexplained electron density had been observed in either site during crystallographic structure determinations. However this does not rule out the possibility of randomly orientated molecules being present in either site. Recently, using the conditions present in milk (pH 6.6 and ionic strength 75mM), BLG has been isolated with 0.5 moles of FFA bound [240]. From an analysis of FFA binding to bovine BLG *in vitro* [176] it has been deduced that the primary binding site is in the external channel (Kd=3μM), whilst several molecules can bind at the weaker secondary binding site which is believed to be the hydrophobic cavity (Kds around 1000μM). As retinol and PNPP have Kd values of 0.02μM [8] and 31μM [164] respectively for their complexes with bovine BLG, it seems likely that they could displace FFAs from the hydrophobic cavity. However the isolation procedures used for bovine BLG cause a dissociation of the protein which is likely to aid the release of any ligand found in the external channel *in vivo*. This implies that both sites could be available for ligand binding studies.

4:2 EXPERIMENTAL DETAILS.

4:2:1 CD AND POLARIMETRY EXPERIMENTAL DETAILS.

The experimental conditions used in the subsequent

CD and polarimetric studies have been described {Chapter 3:2:2 and 3:2:3}. Protein solutions were made up by dissolving sufficient Pentex bovine BLG or pig BLG in 0.1M phosphate buffer so as to obtain A_{280} values between 0.8 and 1.2. The pH values of these solutions were then measured. The following concentrated ligand solutions were made up:

(a) 100mM PNPP in 0.1M phosphate buffer pH 7

(b) 110mM retinol in 95% EtOH

and (c) 3mM biliverdin in 0.1M phosphate buffer pH 9, high concentrations were inhibited by solubility problems. Blanks of each ligand were run, and although none showed any distinct bands, the baselines observed were subsequently subtracted from the appropriate samples. Small aliquots of ligand solution were then added to the protein solutions as required, and the pHs remeasured. The molar ratios of ligand to protein subunit in the samples are listed below:

	Ligand/BLG subunit	%EtOH present
Pig BLG+retinol	1.4	1%
Bovine BLG+retinol	1.4	1%
Bovine BLG+PNPP	1.7	nil
Bovine BLG+biliverdin	0.3	nil

The data obtained were interpreted as described in Chapter 3:2.

4:2:2 TRP FLUORESCENCE EXPERIMENTAL DETAILS.

(a) Solutions.

Stock solutions of protein were made up by dissolving known, weighed amounts of Pentex bovine BLG or pig BLG in 0.1M phosphate buffer at pH 7.3. Small aliquots (about 30μl) of this stock solution were then diluted with 2-3ml of a 0.1M phosphate buffer at the appropriate pH, to give the native protein samples. The

pHs of these were then checked.

Concentrated ligand solutions were made up, such that the addition of a few microlitres of them to the BLG samples produced a 2- to 4-fold excess of ligand, without significantly diluting the protein. Those ligands which were water-soluble were dissolved in 0.1M phosphate buffer at pH 7.3, whereas the remainder were dissolved in 95% EtOH.

The resultant percentage of EtOH had to be kept very low, as alcohol affects both the pH and Trp fluorescence of BLG. An increase of 1% in the EtOH content was shown to raise the apparent pH of the BLG sample by 0.02 pH units, and can be corrected for. A progressive increase in the alcohol content of the BLG sample affects its Trp fluorescence, presumably by altering the conformation of the protein. From preliminary experiments it was shown that the Trp fluorescence of BLG was not significantly altered by 6.5% EtOH under acidic conditions, but was enhanced by only 3.4% EtOH at pHs > 7 {Figure 4:1}. Thus in all subsequent experiments the EtOH content of the BLG samples was kept below 2%.

(b) Fluorimetry conditions.

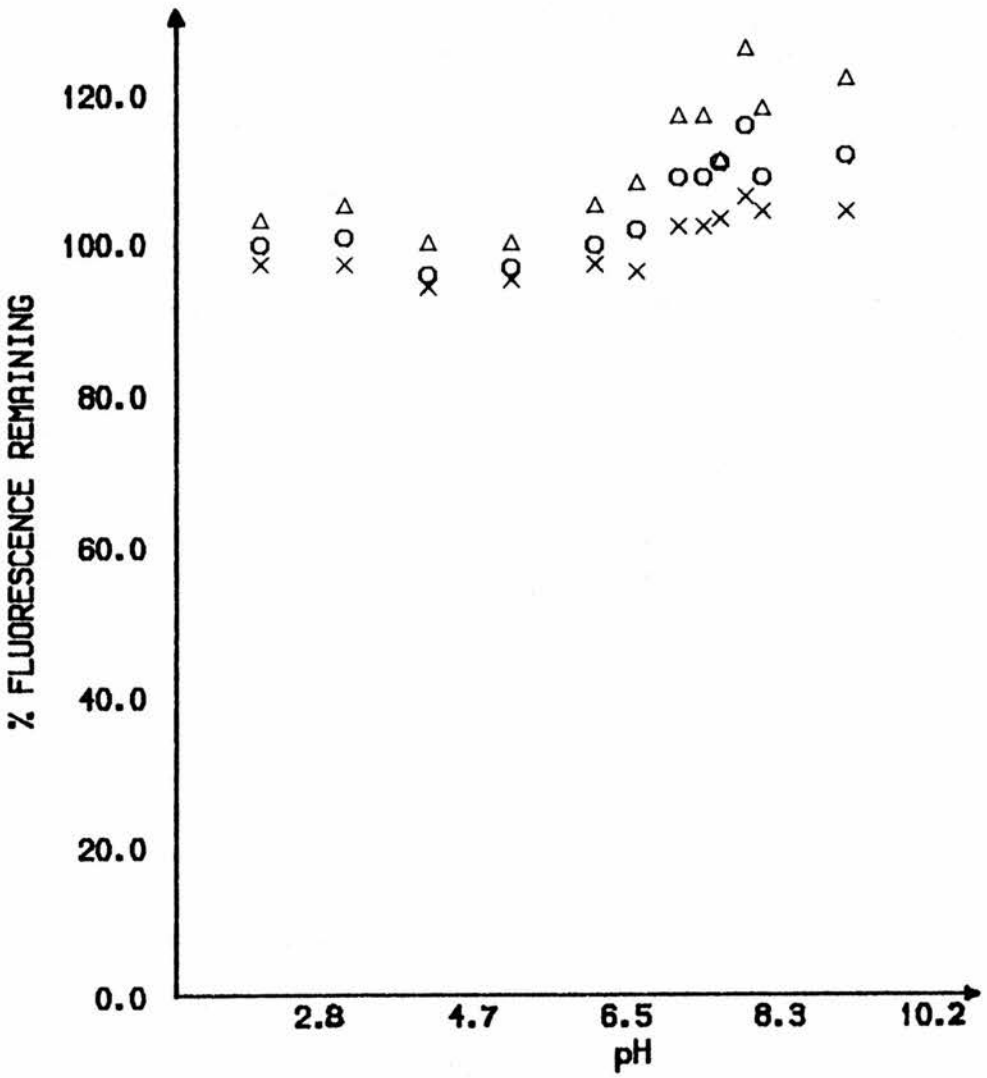
Fluorimetry experiments were carried out on a Perkin-Elmer 3000 fluorescence spectrometer with the samples being placed in a quartz cuvette of pathlength 1cm. The experiments were run, at room temperature and in air, as preliminary studies had indicated that there was negligible quenching of the Trp fluorescence of either pig or bovine BLG by atmospheric oxygen. Both the emission and excitation slits were set at 5nm, and the wavelengths used are given below:

Type of fluorescence.	Excitation λ.	Emission λ.
Tryptophan	295nm	331nm
Retinol	331nm	468nm

Figure 4:1

THE EFFECT OF ETHANOL ON THE TRP FLUORESCENCE OF BOVINE BLG AT VARIOUS pH VALUES.

Excitation wavelength = 295nm
Emission wavelength = 331nm



KEY
X 0·4% EtOH
O 3·4% EtOH
Δ 6·5% EtOH

For Trp fluorescence the excitation wavelength was selected so as to prevent any significant absorption by Tyr and Phe moieties [317], whilst the maximum emission wavelength was determined experimentally. It was found to be 331nm for both pig and bovine BLGs, and independent of pH.

4:3 ANALYSIS OF TRP FLUORESCENCE.

4:3:1 THEORY OF TRP FLUORESCENCE.

The aromatic amino acids, Tyr, Trp and Phe, all absorb UV radiation at 280nm. However at 295nm only Trp still absorbs significantly [317], generating a molecule in an excited π^* electronic state which can fluoresce, its emission wavelength being red-shifted with respect to the excitation wavelength. Both the fluorescence intensity and the maximum emission wavelength are very sensitive to the local environment of the Trp residues. The fluorescence intensity is reduced, when the Trp surroundings become more polar, as the rate of transfer between excited states is enhanced, whilst the maximum emission wavelength is red-shifted.

The fluorescence intensity depends upon the fraction of molecules that become deexcited by fluorescence (the quantum yield), and is quenched when other deexcitation processes occur [318]. These include energy transfer to a neighbouring chromophore in the same molecule which can provide information on the distance between fluorophores, and deexcitation by the addition of a solute. The solute may cause either static or dynamic quenching. The former involves the formation of a complex between the solute and chromophore that predates excitation, which will be instantaneously quenched, thus reducing the concentration of excited molecules and the fluorescence intensity. In

dynamic quenching collisions with the solute provide an additional process for depopulating the excited state, leading to shorter fluorescence lifetimes. This collisional quenching can be represented by the Stern-Volmer equation:

$$Ia / Ip = 1 + K_Q[Q]$$

where Ia and Ip are the fluorescence intensities in the absence and presence of solute respectively, and [Q] is the concentration of the quenching solute. K_Q is the quenching constant, and equal to the product of the fluorescence lifetime and the bimolecular rate constant(k). If $k > 1.4 \times 10^{11}\ M^{-1}sec^{-1}$, the maximum rate constant for diffusion-controlled processes, then binding must have occurred between the solute and chromophore. The dissociation constant for this complex can be estimated from $1/K_Q$, but may be a poor estimate as no correction for collisional quenching was made.

4:3:2 DISTINCTION BETWEEN DIFFERENT TRP RESIDUES.

Iodide acts as a collisional quencher of Trp fluorescence, and hence can provide information on the accessibility of Trp residues. Its failure to quench this fluorescence can be explained by either the buried nature of the Trp residues or the interaction of the iodide anion with charged amino acids in the protein.

The Trp fluorescence of bovine BLG is composed of contributions from both Trp-19 and Trp-61. A 100-fold excess of potassium iodide over bovine BLG failed to quench any of this fluorescence, implying that neither Trp residue was accessible. This can be explained from the structure of BLG-A - Trp-19 is buried at the foot of the hydrophobic cavity whilst Trp-61 is exposed, but near several carboxyl groups (Asp-53, Glu-55, Glu-62, Asp-64 and Glu-65) which could repel the iodide anion. Pig BLG contains only a single Trp residue, Trp-19, and hence all

its Trp fluorescence must be due to this residue. As Trp-19 is conserved in all BLGs and is likely to be important for their function in vivo, its environment should be similar in both pig and bovine BLGs. Thus the additional Trp fluorescence from bovine BLG may tentatively be assigned to Trp-61.

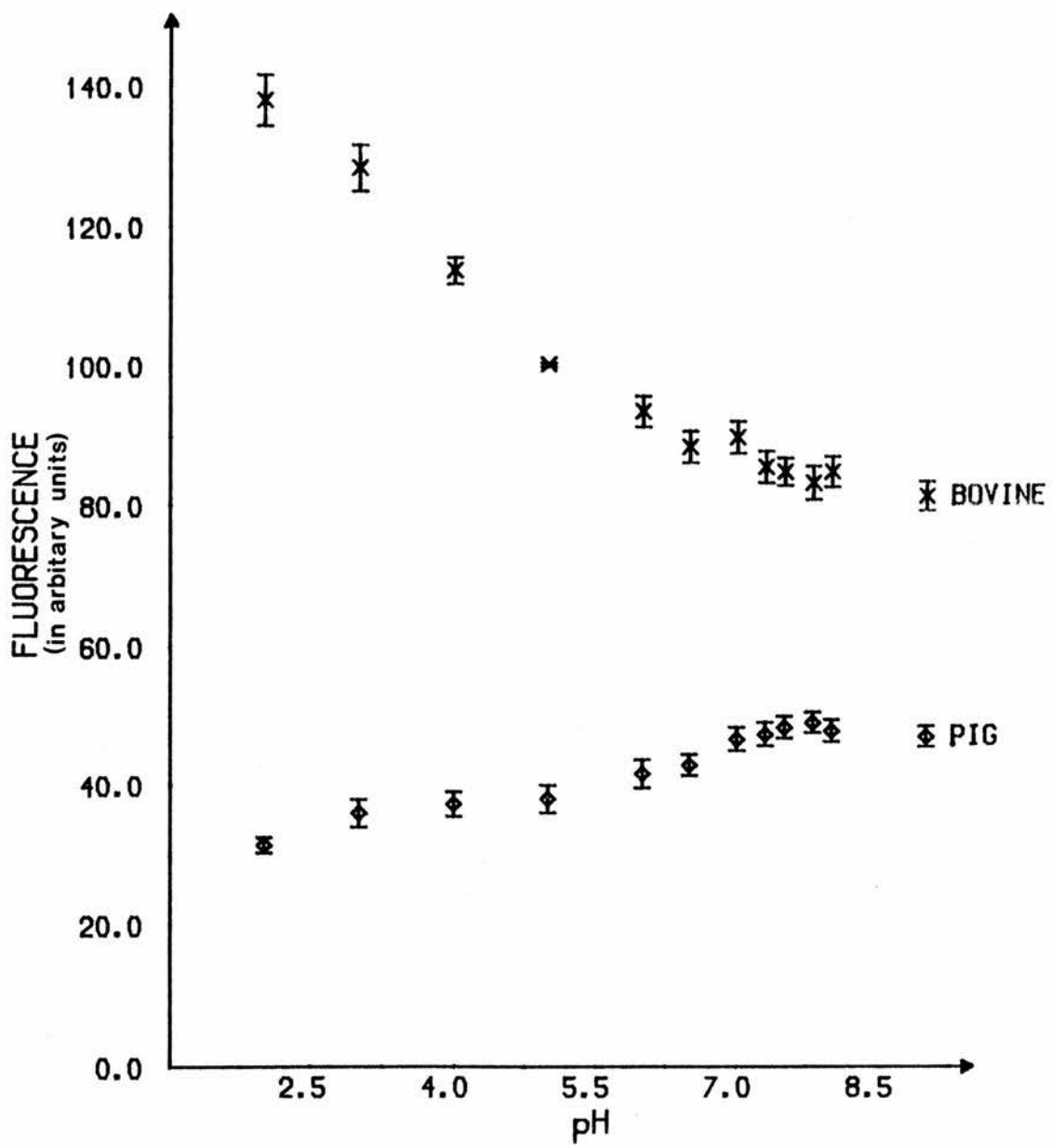
The Trp fluorescences from both bovine and pig BLG, as a function of pH, are given in Figure 4:2. Pig BLG exhibits a constant Trp fluorescence over the pH range 3 to 6 but a slight increase between pH 6 and 8, implying that the Trp-19 environment becomes more hydrophobic. This agrees with the observation made on bovine BLG that the accessibility of Trp-19 to solvent is reduced in lattice Y (pH 7.8) compared to lattice X (pH 6.5) {Chapter 3:3:4}. For bovine BLG the variation in Trp fluorescence is considerably different due to the contribution from Trp-61. The large decrease in fluorescence over the pH range 2.5 to 7 may be due to the creation of a more polar environment by the increased ionization of carboxyl side chains. The typical pK value for both Asp and Glu residues in a protein is 4.5, which corresponds to the midpoint of this large decrease in fluorescence. The ratio of the Trp fluorescence of pig BLG:bovine BLG at pH 7 correlates excellently with the ratio of the absorption coefficient values of 5.65 and 9.70 obtained from the A_{280} measurements on 1% solutions of pig and bovine BLGs solutions respectively. This suggests that the same proportion of the radiation absorbed by each protein is emitted as fluorescence, thus eliminating the possibility of significant energy transfer from Trp-19 to Tyr-20, as the latter is absent in pig BLG.

4:3:3 INTERPRETATION OF LIGAND BINDING DATA.

The addition of any ligand to a solution of BLG may

Figure 4:2

THE TRP FLUORESCENCE OF PIG AND BOVINE BLGS OVER THE pH RANGE 2 TO 9.



affect the Trp fluorescence of the protein in several ways. There may be no effect if the Trp residues are inaccessible, as observed with the addition of iodide; or a decrease in fluorescence if collisional quenching occurs, or the ligand binds to BLG and alters either the conformation or environment of its Trp residues. Data were collected, at any specific pH, by measuring the Trp fluorescence from four types of sample:

BLG in buffer	Fluorescence = A
Buffer	Fluorescence = B
Ligand in buffer	Fluorescence = C
BLG + ligand in buffer	Fluorescence = D

and then presented in the following manner:

Fluorescence of native BLG = (A - B)

Fluorescence of BLG with ligand present = (D - C)

hence % Fluorescence remaining = {(D - C)/(A - B)} x 100

The error bars applied to these values in subsequent figures represent the standard error of the mean calculated from at least three measurements. However this approach makes several assumptions:

(i) the ligand is present in excess, so that the amount of free protein is insignificant. This could be valid, as it is known that a BLG+retinol complex has a very low dissociation constant (Kd = 0.02μM) [153];

(ii) the volume of the ligand solution should be small compared to that of the protein solution so that the dilution of BLG is negligible. This was achieved by using high concentration ligand solutions such that the dilution of the BLG was always < 0.5%;

(iii) that the 'ligand+buffer' is an adequate blank for the BLG+ligand sample. This is a poor approximation for the retinoids which not only absorb at the emission wavelength of the Trp fluorescence, but could also provide collisional quenching. A blank containing N-acetyl-L-tryptophanamide (which does not bind to retinol) and retinol showed a decrease of about 15% of that

observed when the true protein was used [297]. With the other ligands to be studied, which do not absorb at 330nm, the decrease should be smaller and due to just collisional quenching. Decreases of below 10% in Trp fluorescence upon ligand binding to BLG were subsequently taken as representing no significant static quenching.

4:4 BINDING OF VARIOUS LIGANDS TO BLG.

4:4:1 VARIETY OF LIGANDS TO BE STUDIED.

Bovine BLG is able to bind a wide range of small molecules {Chapter 1:7}. Many of those eg. free fatty acids, alkanes and SDS, are believed to bind at the same site, implying that this site may be fairly specific for long chain hydrocarbons. This site is likely to be the external hydrophobic channel as both it, and the association constant for the binding of FFA to BLG, vary over the Tanford transition. Retinol has also been found here in the crystallographic structure of BLG lattice Z+retinol. The binding site for other ligands is less definite, but could be within the hydrophobic cavity. Iodobenzene has been detected here by crystallographic studies, whilst retinol has been modelled into the cavity (where it binds in RBP [215]), and PNPP may also bind in this location, as it has been reported to affect the Trp fluorescence of bovine BLG [164]. However the differing nature of these compounds, together with the large volume of the hydrophobic cavity (360Å^3) suggests that this binding site may be available to a wide range of ligands, and hence non-specific. As Trp-19 is located at the foot of the calyx, the binding of a variety of ligands to both pig and bovine BLG was investigated by Trp fluorescence.

The ligands used in this study were chosen for a number of reasons, and the structures of some of them are

given in Figure 4:3. The structural homology between BLG, INCYN, and RBP, indicated that the natural ligands of the latter two might bind to BLG. Thus biliverdin, retinol and two other retinoids were selected. Other vitamins were also included: biotin(vitamin H) as its binding protein streptavidin has a β-barrel structure [238], α-tocopherol(vitamin E) since it is a fat-soluble vitamin like retinol, and vitamins K1 and K2. Ligands of some of the other superfamily proteins were also chosen eg. pyrazines and progesterone. Various p-nitrophenyl compounds, which have been reported to bind to bovine BLG and affect its Trp fluorescence, [164] were also examined.

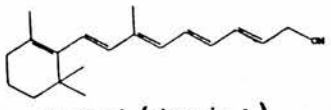
4:4:2 LIGAND BINDING TO PIG AND BOVINE BLGS.

The Trp fluorescence observed, after the addition of approximately a 3-fold excess of ligand to the BLG solution, was expressed as a percentage of the fluorescence from the original protein solution (% fluorescence remaining). The results obtained from the addition of various ligands to both bovine and pig BLGs are tabulated in Figures 4:4 and 4:5.

There appears to be no simple correlation between the decrease in fluorescence observed from bovine BLG with retinol, PNPP, p-nitrophenyl acetate or p-nitrophenyl-β-glucuronide, and the dissociation constants of these complexes (0.02, 31, 33 and 64μM respectively [153,164]). This was expected, as the fluorescence observed is the result of contributions from Trp-19, Trp-61 and solute quenching, which have not been resolved. A comparison of the effect of a specific ligand on both pig and bovine BLGs may give an indication of the Trp-61 contribution (assuming that the only Trp in pig BLG, Trp-19, is in a similar location in bovine BLG); whilst the titration of BLG with the ligand should reveal the

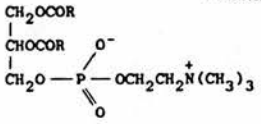
Figure 4:3

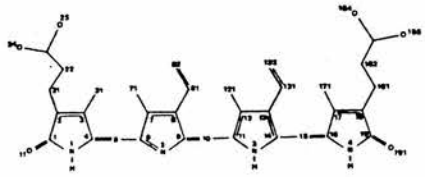
THE STRUCTURES OF SOME POSSIBLE LIGANDS FOR BLG.



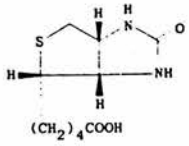
Retinol (vitamin A)

Lecithin



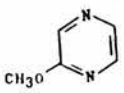


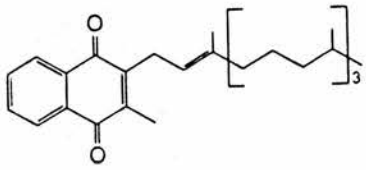
Biliverdin IXγ



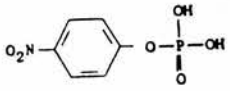
Biotin

2-Methoxypyrazine

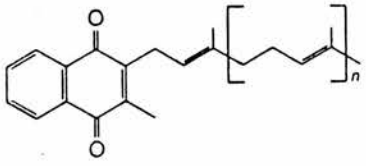




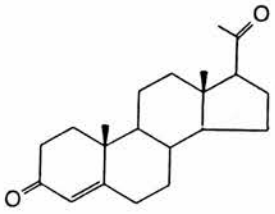
Phylloquinone
(vitamin k_1)



p-Nitrophenyl Phosphate



Menaquinone
(vitamin k_2)



Progesterone

Figure 4:4

CHANGES IN THE TRP FLUORESCENCE OF BOVINE BLG UPON THE ADDITION OF VARIOUS LIGANDS.

Ligand / Solvent	pH:	% Trp fluorescence remaining 6.0	6.5	7.0	7.3	7.5	7.8	8.0
Retinol/ EtOH		49.8	48.9	30.5	26.0	20.1	15.1	16.6
Retinyl acetate / EtOH		77.8	77.2	68.5	74.9	73.4	75.6	68.0
Retinyl palmitate / EtOH		78.0	80.2	79.1	79.2	72.4	76.3	79.4
p-Nitrophenol / 0.1M P pH 7.3		79.8	82.3	83.4	88.4	90.0	92.6	94.1
p-Nitrophenyl phosphate / 0.1M P pH 7.3		78.5	77.8	78.7	78.5	79.5	79.4	78.5
p-Nitrophenyl acetate / 0.1M P pH 7.3		87.1	89.5	95.9	90.6	91.4	93.6	95.2
p-Nitrophenyl-β-glucuronide / 0.1M P pH 7.3		76.8	77.9	77.2	78.6	78.3	80.7	80.5
Biliverdin / 0.1M phosphate pH 9		61.8	64.5	67.1	64.4	61.9	64.8	62.0
2-methoxypyrazine / water		90.2	90.5	91.5	91.3	90.4	88.8	91.3
Pyrazine / water		96.4	97.7	97.2	98.6	96.7	97.2	95.3
Amyl acetate / water		96.8	96.6	97.5	95.0	98.6	97.1	98.6
Progesterone / EtOH		96.6	97.8	98.9	98.4	98.3	96.4	96.9
α-lecithin / EtOH		96.2	96.9	98.1	98.5	101.2	101.3	101.6
Biotin / 0.1M phosphate pH 8		95.7	97.3	98.6	98.9	99.9	100.0	100.3
α-tocopherol / EtOH		103.4	99.9	105.2	95.9	92.9	98.5	95.7
Vitamin K1 / EtOH		-	-	99.9	-	-	-	-
Vitamin K2 / EtOH		-	-	94.3	-	-	-	-
SDS / water		103.1	103.4	104.9	105.6	106.5	108.0	109.9
L-dopa / 0.1M phosphate pH 2		88.1	92.9	90.9	89.4	91.5	93.4	92.5

Figure 4:5

CHANGES IN THE TRP FLUORESCENCE OF PIG BLG UPON THE ADDITION OF VARIOUS LIGANDS.

Ligand / Solvent	pH :	% Trp fluorescence remaining 6.0	6.5	7.0	7.3	7.5	7.8	8.0
Retinol / EtOH		53.2	49.3	42.7	43.7	42.4	43.3	41.2
Retinyl acetate / EtOH		-	86.0	90.6	88.8	81.8	84.2	82.0
Retinyl palmitate / EtOH		61.2	78.4	73.3	74.3	76.7	65.8	72.9
p-Nitrophenyl phosphate / 0.1M P pH 7.3		83.0	83.2	84.8	86.4	81.7	81.8	82.0
p-Nitrophenyl acetate / 0.1M P pH 7.3		90.0	93.0	93.9	92.3	97.3	87.2	91.1
p-Nitrophenyl-β-glucuronide / 0.1M P pH 7.3		86.0	85.1	85.4	85.0	89.8	86.9	87.9
Biliverdin / 0.1M phosphate pH 9		41.5	37.4	36.0	47.5	37.8	38.8	40.7
2-methoxypyrazine / water		92.9	91.3	93.8	92.9	95.2	89.7	93.0
Pyrazine / water		93.9	96.1	96.4	93.9	97.4	96.6	93.3
Vitamin K1 / EtOH		-	-	81.9	-	-	-	-
Vitamin K2 / EtOH		-	-	67.1	-	-	-	-
SDS / water		97.5	97.7	97.6	97.2	98.6	95.5	96.5
L-dopa / 0.1M phosphate pH 2		73.9	82.4	83.8	82.0	81.2	89.1	87.8

importance of solute quenching.

Changes of less than 10% in the '% fluorescence remaining' when a 3-fold excess of ligand was added to BLG, are unlikely to indicate specific binding between the ligand and Trp residues in the protein. Hence it was proposed that neither pyrazine nor 2-methoxypyrazine, ligands of the odorant binding proteins, bind near the Trp residues of BLG. This seems reasonable as it has been reported that 2-isobutyl-3-methoxypyrazine cannot bind to bovine BLG [247]. Progesterone (which binds to A1AG), and α-lecithin (which forms a complex with APOD) are also unlikely to bind near the Trp residues in BLG. This suggests that the hydrophobic cavity, with Trp-19 at its foot, is not a non-specific binding site for the ligands of other members of the BLG superfamily. The binding of retinol to BLG fits this idea, as, to date, it has only been found in the external channel [59].

From these Trp fluorescence studies on the binding of ligands to pig and bovine BLGs three ligands were selected for more detailed analyses:

(A) retinol was chosen because its addition to BLG caused a large variation in the '% fluorescence remaining' which was dependent upon pH, unlike that observed with any other ligand,

(B) biliverdin, as its presence produced the largest constant decrease in Trp fluorescence of BLG over the pH range studied, and

(C) p-nitrophenyl phosphate, which, due its solubility in water, should be easy to get in the same medium as BLG, and may bind at a different, non-hydrophobic site.

4:5 PRELIMINARY BINDING STUDIES WITH RETINOL, PNPP AND BILIVERDIN.

Preliminary studies were carried out to determine

the effect of the ligands, retinol, PNPP, and biliverdin, on the emission wavelength of the Trp fluorescence of BLG. The quenching of this fluorescence by these ligands was examined using Stern-Volmer plots.

4:5:1 VARIATION IN TRP FLUORESCENCE EMISSION WAVELENGTH.

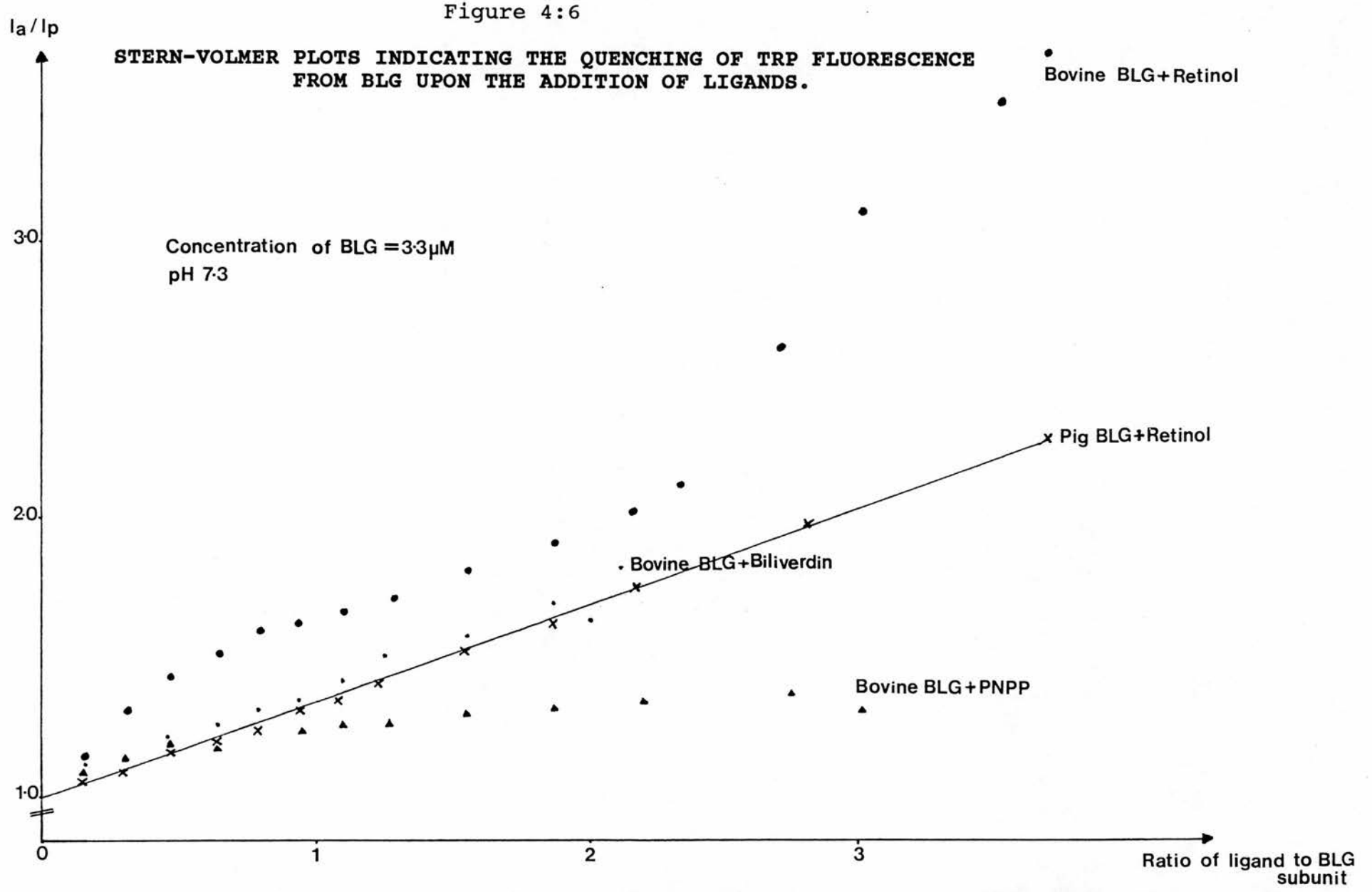
The Trp fluorescence of 3.3μM solutions of bovine BLG (pHs 6, 7.3, and 8), with and without a 3-fold excess of ligand, were measured across the emission wavelength range 310-360nm. Upon addition of each of the three ligands, retinol, PNPP, and biliverdin, the maximum emission wavelength remained constant at around 330nm. This implied that none of these compounds altered the hydrophobicity of the Trp environments in BLG.

4:5:2 ORIGIN OF TRP FLUORESCENCE QUENCHING.

The Trp fluorescence of bovine BLG originates from residues Trp-19 and Trp-61, whereas that of pig BLG only comes from Trp-19. In both cases the addition of the above ligands caused a decrease in the protein fluorescence, and Stern-Volmer plots were used to analyse the quenching {Figure 4:6}.

The addition of retinol to pig BLG at pH 7.3 caused a decrease in Trp fluorescence that produced a linear Stern-Volmer plot. The gradient of this line revealed that the dynamic(collisional) quenching constant was $101,000M^{-1}$. By assuming that pig BLG has the same fluorescent lifetime as that of bovine BLG (1.4 nanoseconds) [153], the bimolecular rate constant for quenching was calculated to be $7.2x10^{13}M^{-1}s^{-1}$, which is greater than the maximum rate constants for diffusion controlled processes. This implies that retinol binds to pig BLG with a Kd value of around 10μM, which is greater than the 0.02μM reported for the bovine BLG+retinol

Figure 4:6


STERN-VOLMER PLOTS INDICATING THE QUENCHING OF TRP FLUORESCENCE FROM BLG UPON THE ADDITION OF LIGANDS.
Ia / Ip
Concentration of BLG =3·3μM
pH 7·3
Bovine BLG+Retinol
Pig BLG+Retinol
Bovine BLG+Biliverdin
Bovine BLG+PNPP
3·0
2·0
1·0
0
1
2
3
Ratio of ligand to BLG subunit

complex [153].

The Stern-Volmer plot obtained upon the addition of retinol to bovine BLG was more difficult to interpret. The non-linearity observed up to 2 moles of retinol per subunit of BLG can be represented by two lines with K_Q values of $212,000M^{-1}$ and $100,000M^{-1}$. As the latter is similar to that observed for pig BLG+retinol, this implies that the initial quenching of the bovine BLG Trp fluorescence involves both Trp-61 and Trp-19. The addition of more retinol caused an upward curve in the Stern-Volmer plot, indicating additional binding. This would fit with the idea that retinol initially binds in the external channel which is near Trp-19 (Trp-19 -> Cys-121 is 9.1A), and then in the hydrophobic cavity, where its much closer approach to Trp-19 causes the large reduction in fluorescence observed.

A previous report has revealed that the addition of PNPP to bovine BLG produced a linear Stern-Volmer plot with a K_Q of $28,800M^{-1}$ [164]. However no data were recorded on low ratios of PNPP:BLG. From our plot linearity was observed up to a 1:1 ratio of ligand:BLG with a gradient of $84,800M^{-1}$ probably due to the quenching of both Trp-61 and Trp-19 fluorescence. Using higher PNPP concentrations caused a reduction in the gradient, producing a line with a K_Q value of $27,500M^{-1}$ which is in excellent agreement with that published [164]. This incomplete quenching suggests that Trp-61, which is situated near several ionized, carboxyl side chains, may become partially inaccessible to PNPP because of the latter's charged phosphate moiety.

The addition of biliverdin to bovine BLG produced a linear Stern-Volmer plot. This suggests that either both Trp residues are quenched similarly, or that only one is quenched. The latter seems more likely as the size of biliverdin could prevent it entering the hydrophobic cavity and quenching fluorescence from Trp-19. Its

location at the entrance to the cavity would correspond with its position in INCYN and BBP [222,223], and enable it to quench Trp-61 fluorescence. The quenching constant is 110,000M^{-1}, which indicates that the dissociation constant for this complex is about 9μM.

4:6 RETINOL BINDING TO BLG.

4:6:1 CURRENT KNOWLEDGE.

From the structural homology between bovine BLG and RBP, retinol has been proposed as the ligand for BLG *in vivo* [8]. The complex formed between bovine BLG and retinol in solution has been characterized [153]. The use of ^{3}H-retinol indicated that one mole of retinol bound per subunit of BLG, and this was confirmed by a Scatchard plot based on the induced CD signal at 330nm, which gave a Kd value of 0.02μM. Retinol fluorescence studies inferred that the binding was not pH-dependent between pH 2 and 7.5, although measurements were only reported at these two values. Absorption spectroscopy revealed a red shift upon binding, implying that the binding site was hydrophobic [153].

The detailed nature of the binding site has only been partially elucidated. HNBB modification of all four Trp residues per dimer of BLG appeared to alter the interaction between BLG and retinol by removing the vibrational fine structure of retinol and blue-shifting the absorption maximum to that of free retinol. However fluorescence polarization indicated that retinol was still bound. As HNBB is also known to modify free thiols, the above observations could be explained if retinol binds in the external channel which involves Cys-121, and is then released upon modification of the thiol, and then binds within the hydrophobic cavity.

The structural homology between RBP and BLG suggested such a binding site for retinol within the hydrophobic cavity. However, the lack of a collar of Phe residues part way down the cavity in BLG should enable retinol to go deeper into the cavity than it does in the RBP+retinol complex [215]. Retinol has been modelled into the hydrophobic cavity of BLG and then energy minimized, confirming a plausible position for it deeper down than in RBP. The shorter distance between retinol and Trp-19 in BLG could suggest stronger binding. This is supported by the smaller Kd value observed for the BLG+retinol complex compared to that for the RBP+retinol complex (0.02 compared to 0.2μM); but does not rule out the idea that binding could occur initially at the strongly binding external channel, before binding in the cavity, and that the Kd obtained is for the initial binding. BLG is known to bind retinol in the hydrophobic channel near Cys-121 and the α-helix [59], and as the α-helix is believed to be linked with the Tanford transition {Chapter 3:3:1} this binding site could undergo a conformational change over the pH range 6 to 8, enabling a ligand to be released to its receptor.

To try and resolve the anomaly between the two possible binding sites for retinol in BLG, and provide more information on the residues in the binding sites, and their possible association with the Tanford transition, a variety of solution experiments were undertaken.

4:6:2 RETINOL BINDING TO PIG BLG.

The titration of pig BLG (pH 7.3) with retinol was monitored by Trp fluorescence, and indicated that a 1:1 complex formed. The linear Stern-Volmer plot obtained {Figure 4:6} confirmed binding near the single chromophore of pig BLG, Trp-19, and showed that the

complex had a Kd around 10μM {Chapter 4:5:2}. However the closeness of Trp-19 to residues in the external channel indicates that binding could be at either site.

The reduction in Trp fluorescence, as a function of pH, upon the addition of a 3-fold excess of retinol to pig BLG was examined {Figure 4:7}. The '% fluorescence remaining' stayed constant around 55% between pH 2 and 6, but decreased to 41% as the pH was raised to 8 - suggesting that the enhanced quenching might be linked to the Tanford transition. This conformational change could increase the exposure of Trp-19 to solvent and enhance the collisional quenching, or strengthen the binding of retinol to pig BLG and increase the static quenching. The former appears unlikely, as in bovine BLG Trp-19 is known to become more buried at higher pH (Chapter 3:3:4). This implies that the conformational change enhances the binding of retinol to pig BLG - the same trend is observed when palmitate binds in the external channel of bovine BLG [176].

The position of retinol in pig BLG is unknown. As yet suitable crystals of this protein have not been grown, thus preventing the use of soaking experiments and x-ray crystallography to identify the location of the ligand. Near UV CD spectra showed that the addition of retinol to pig BLG did not affect the ellipticities of the signals at 285 and 293nm (which are due to Tyr and Trp residues respectively) at pH 6, but reduced their strength at pH 8 {Figure 4:8}. This seems to confirm that retinol cannot be situated very near Trp-19 at pH 6, which seems surprising if the cavity is the binding site, as pig BLG does not possess the collar of Phe residues found in RBP, but supports the idea of binding in the external channel. However at pH 8 the retinol could get closer to Trp-19 as it becomes more firmly bound into the channel, and this would explain the increased quenching of Trp fluorescence as the pH increases across the

Figure 4:7

THE QUENCHING OF TRP FLUORESCENCE FROM PIG AND BOVINE BLGS UPON THE ADDITION OF RETINOL.

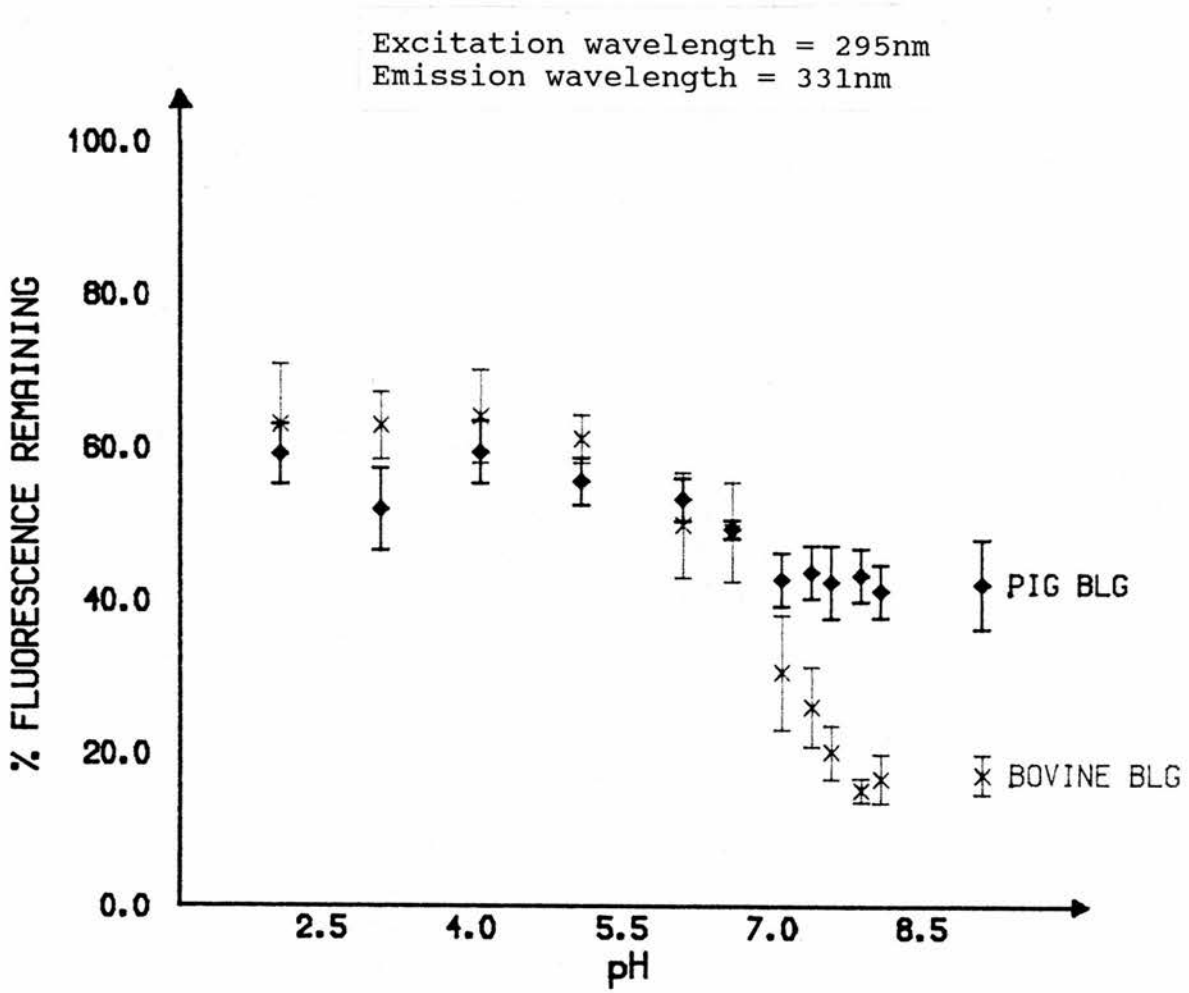


Figure 4:8

THE CHANGES OBSERVED IN THE NEAR UV CD SPECTRA OF PIG AND BOVINE BLGS, AT pH VALUES 6 AND 8, UPON THE ADDITION OF LIGANDS.

Solution	pH	[X](MRW) 267nm	285nm	293nm
Pig BLG	6	-	-97.50	-85.00
Pig BLG+retinol	6	-	-96.64	-85.71
Pig BLG	8	-	-103.65	-91.67
Pig BLG+retinol	8	-	-98.46	-85.71
Bovine BLG	6	-45.51	-71.79	-89.10
Bovine BLG+retinol	6	-39.57	-70.21	-86.17
Bovine BLG+PNPP	6	-42.71	-73.98	-90.00
Bovine BLG+biliverdin	6	-44.22	-79.59	-93.88
Bovine BLG	8	-33.83	-57.45	-65.74
Bovine BLG+PNPP	8	-31.27	-57.20	-68.64
Bovine BLG+biliverdin	8	-54.04	-104.61	-122.55

Tanford transition.

4:6:3 RETINOL BINDING TO BOVINE BLG.

The titration of bovine BLG with retinol was recorded by Trp fluorescence, and revealed a non-linear Stern-Volmer plot, indicating that the two Trp residues, Trp-19 and Trp-61, did not quench similarly. At higher concentrations of retinol the Stern-Volmer plot curves upwards indicating additional binding, perhaps in the internal hydrophobic cavity. Pig BLG, however, showed no additional binding - this was surprising, but plausible, considering the large number of amino acid substitutions between pig and bovine. One or more of these could block the cavity, and prevent the molecules approaching near to Trp-19 at the calyx.

The '% fluorescence remaining' upon the addition of a 3-fold excess of retinol to bovine BLG was examined over the pH range 2 to 9 {Figure 4:7}. The value remained constant around 60% between pH 2 and 6 which was very similar to that observed from pig BLG+retinol, and could indicate binding at a conserved site in both proteins. If this site is the hydrophobic channel then the substituted residues, eg. Tyr-102 and Cys-121, cannot be critical for the binding interaction. Upon increasing the pH across the Tanford transition the quenching of the fluorescence is enhanced, more so in the case of bovine than pig BLG. It may be coincidence, but the reduction in '% fluorescence remaining' between pH 6 and 8 is 34% for bovine compared to 12% for pig, the latter being about a third of the former - the same ratio as that observed in the change of optical rotation over the Tanford transition for these two proteins. The enhanced quenching across this transition is likely to be due the conformational change modifying the binding site, and strengthening the binding interaction.

Near UV CD spectroscopy was carried out on solutions of bovine BLG+retinol at pHs 6 and 8 to try and identify which residues were affected by the ligand {Figure 4:8}. At pH 6 the presence of retinol affected the ellipticities of the bands at 267nm (Phe) and 293nm (Trp), suggesting the retinol binds near Phe-136 in the external channel, and not far from Trp-19. The spectrum obtained at pH 8 was disregarded as the protein had partially denatured overnight, and time prevented us repeating this experiment.

The association of retinol with bovine BLG has also been monitored by retinol fluorescence [153,161]. The latter report revealed that the complex formed was probably 1:1 retinol:BLG subunit, but did not rule out a 2:1 complex. The binding of retinol to the bovine protein at pH 7.4 enhanced the fluorescence yield of the ligand by about 3-fold [161], which agrees well the 3.5-fold fluorescence enhancement obtained by ourselves. In addition the enhancement was recorded at other pH values across the Tanford transition, and revealed a linear increase from pH 6 to 8, considerably greater than that obtained with pig BLG {Figure 4:9}. The pH-dependence of this enhancement seems to confirm binding in the external channel, whilst the difference between pig and bovine BLGs should correlate with the substituted residues in this region. The binding of retinyl acetate to bovine BLG causes an enhancement in retinol fluorescence of only 10%, suggesting that the length of this derivative may be too long to bind in the external channel.

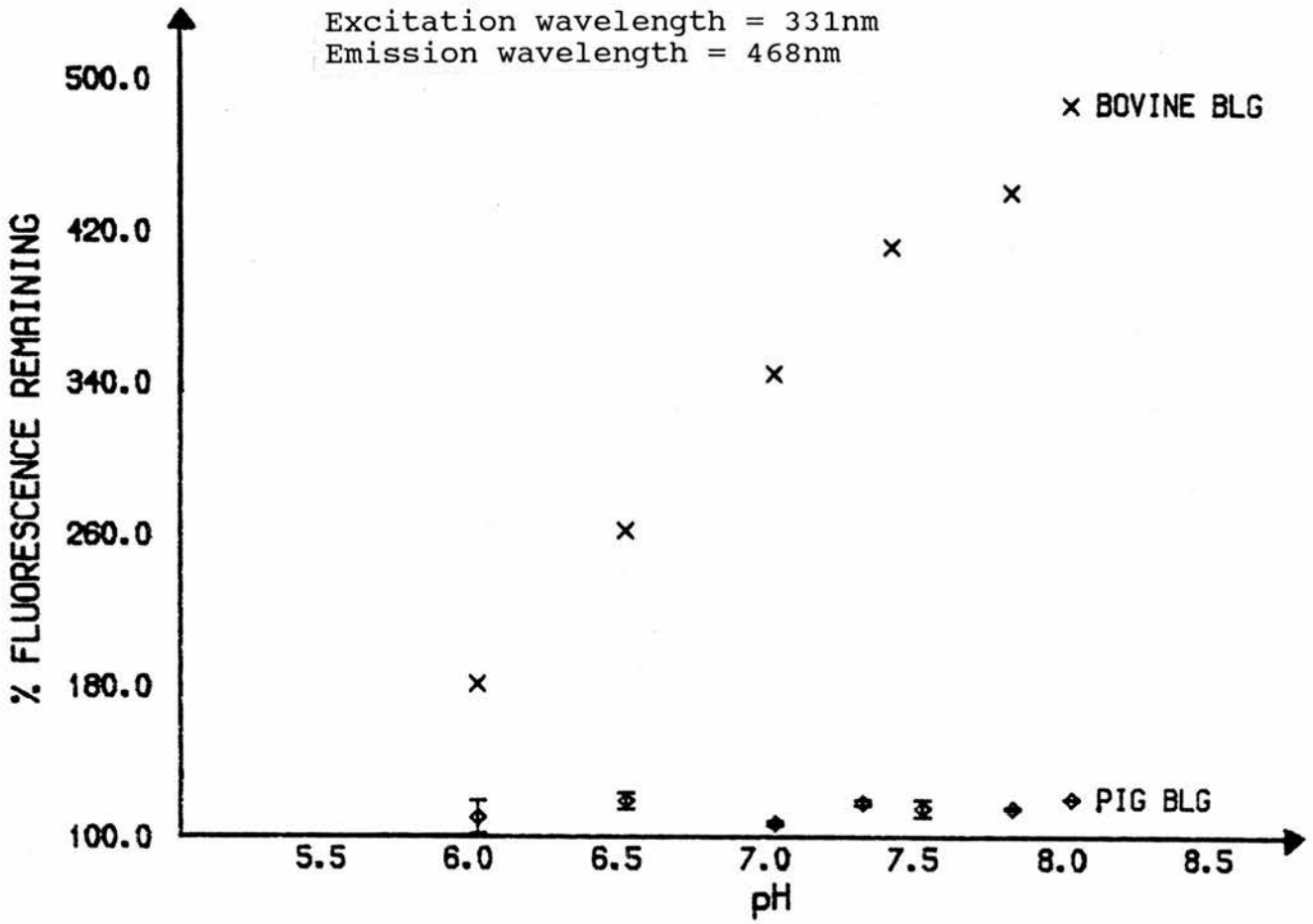
4:7 P-NITROPHENYL PHOSPHATE BINDING TO BLG.

4:7:1 CURRENT KNOWLEDGE.

The binding of several p-nitrophenyl compounds to

Figure 4:9

THE ENHANCEMENT OF RETINOL FLUORESCENCE OVER THE pH RANGE 6 TO 8 UPON THE ADDITION OF PIG AND BOVINE BLGS.



BLG in solution has been studied, and the dissociation constants of these complexes reported [164]. The association of p-nitrophenyl phosphate(PNPP) with bovine BLG was examined in further detail.

Gel filtration at pH 6 revealed that one mole of PNPP bound per subunit of BLG with a Kd of 60μM; whilst fluorescence quenching (excitation wavelength 280nm) confirmed the presence of a complex, and yielded Kd values of 31, 63, and 70μM for the bovine BLG variants A, B, and C respectively [164]. This implies that one or both of the substitutions between variants A and B, Asp-64 -> Gly and Val-118 -> Ala, may be linked with the binding interaction. For BLG-A the Kd value was independent of pH over the range 4 to 7.5, and unaltered if SDS had been bound to the BLG prior to the addition of PNPP [164]. Thus it seems likely that the binding site is hydrophobic, and distinct from the external hydrophobic channel where SDS is believed to bind. CD spectra of bovine BLG at pH 6, in the presence and absence of PNPP, implicated Phe and Trp residues in the association of BLG with this ligand.

Crystallographic studies were carried out on crystals of BLG lattice Y soaked in PNPP, to determine the location of the ligand in the complex, and support the above observations. Additional solution studies were also carried out to investigate the effect of the Tanford transition on this binding.

4:7:2 CRYSTALLOGRAPHY OF THE BLG+PNPP COMPLEX.

(a) Data collection.

Crystals of bovine BLG lattice Y were transferred from their growth medium to mother liquor (2.8M ammonium sulphate + 0.1M phosphate buffer at pH 7.7). A small

aliquot of a concentrated solution of PNPP in 0.1M phosphate buffer at pH 7.5 was added, such that the mixture became 10mM in PNPP. The crystals were allowed to soak for 72 hours at room temperature.

The soaked crystals, which appeared slightly yellow, were mounted in a Lindemann tube. The yellow tinge was believed to be due to the presence of p-nitrophenol (caused by the partial hydrolysis of the PNPP), but did indicate that a p-nitrophenyl ligand had been absorbed into the crystals.

Data were collected, up to 2.8A resolution, from two crystals, using a Siemens-Stöe AED-2 diffractometer with copper radiation (λ=1.5418A). The positions of ten low resolution reflections were recorded from photographic film, and used to determine the orientation matrix, and subsequently check the settings throughout data collection. Profiles were then obtained from scans of about 10 strong reflections, and used to measure the integrated intensities of all the reflections detected. Empirical absorption corrections, derived from psi scans were applied prior to data processing.

(b) Data processing.

For each crystal, A, and B, the data collected were converted from I, sigI to F, sigF using SHELX76, and then equivalent reflections were merged. The merging details are given in Figure 4:10, and indicate that the first dataset appeared reasonable (R_{merge} is 0.038), whilst the second was bad (R_{merge} is 0.122), probably due to the crystal being non-isomorphous. The bad dataset was discarded, and the other concatenated on to those reflections (with their phases) to 2.8A resolution which were present in our resultant native BLG lattice Y dataset.

The data from this BLG+PNPP crystal were then scaled

Figure 4:10

SUMMARY OF THE CRYSTALLOGRAPHIC DATA OBTAINED FROM CRYSTALS OF A BOVINE BLG+PNPP COMPLEX.

	Crystal A	Crystal B
No. measurements	6691	4386
No. acceptable measurements	5664	4084
No. unique measurements	5309	3137
No. used in merging	355	947
R_{merge} (to 2.8A resolution)	0.038	0.122

The data from crystal A were scaled to those present in the resultant native BLG lattice Y dataset.

No. common reflections = 3654

R_{scale} to 2.8A resolution = 17.1%

to the native dataset using the Fox and Holmes scaling procedure [292]. The scale factor was determined from the ratio of the mean intensities of the datasets, but was not refined, and included no exponential terms. A summary of the scaling is given {Figure 4:10}.

(c) Generation of an electron density map.

Using the data from the BLG+PNPP crystal which had been scaled to the native BLG data a 2.8A electron density difference map was generated with coefficients (F_A-F) where F_A is the structure factor for a reflection in the BLG+PNPP dataset. The map was generated over the region:

$$-10/64 < X < 45/64$$
$$-24/80 < Y < 48/80$$
$$-30/96 < Z < 40/96.$$

This map was then displayed on an Evans & Sutherland PS300 using the molecular graphics program FRODO. The electron density levels available ranged from -400 to 365 - by setting the contour level at 350 it was possible to display only the three strongest electron density regions of the map. One appeared to be in an adjacent subunit, one was near Ala-143, and the third was within the β-barrel near Asp-96 and Trp-19. It was approximately 6A from Asp-96, and 5A from Trp-19 {Figure 4:11}. This seems to support the view that a p-nitrophenyl compound can bind within the hydrophobic cavity of BLG, but it is more likely to be the p-nitrophenol formed from the hydrolysis of PNPP than PNPP, as the latter has a charged phosphate moiety which would cause unfavourable interactions within the β-barrel.

4:7:3 SOLUTION STUDIES ON PNPP BINDING TO BLG.

The titration of bovine BLG with PNPP has been

Figure 4:11

PHOTOGRAPH TO ILLUSTRATE THE STRONGEST AREAS OF ELECTRON DENSITY IN THE 2.8A DIFFERENCE MAP OF (BLG+PNPP - BLG).

The main area of density lies within the hydrophobic cavity, near residues Trp-19, Asp-96 and Lys-101. Its distance from Trp-19 is about 5A.

carried out at two pH values, and monitored by Trp fluorescence. The Stern-Volmer plot published was obtained at pH 6 and gave a Kd value around 40μM, whilst that determined by ourselves at pH 7.3 yielded a Kd of 36μM. This may indicate that the strength of binding is independent of pH.

The '% fluorescence remaining ' upon the addition of a 3-fold excess of PNPP to pig and bovine BLG was similar for both proteins, at around 80% {Figure 4:12}. This seems to support the idea that the binding site is conserved in both proteins, whilst the lack of variation in the decrease of Trp fluorescence with pH confirms the hydrophobic nature of the binding site. These observations, together with the crystallographic data, imply that the hydrophobic cavity is a conserved binding site in BLGs and may be the site required for the true function of this protein. Binding of PNPP at this site appears to have no effect on the Tanford transition, as detected by polarimetry {Figure 4:13}.

The position of PNPP in the β-barrel of bovine BLG has been detected by crystallography. Solution studies using near UV CD were carried out to determine the residues involved in the binding. At pH 6 the near UV CD spectrum appeared to show a slight variation in the strength of the band at 267nm (Phe) upon the addition of PNPP to bovine BLG {Figure 4:8}. This implicates Phe residues in the binding site (possible Phe-82 and Phe-105 which line the β-barrel), which agrees with the conclusion drawn by Farrell et al. [164]. However it should be noted that the variations observed are different in their spectra. From a comparison of three native BLG near UV CD spectra [164,175, and ours] it appears that the relative intensities of the bands at 267nm, 285nm and 293nm are different in Farrell's spectrum compared to the other two, suggesting that their protein may not be in the native conformation.

Figure 4:12

THE QUENCHING OF TRP FLUORESCENCE FROM PIG AND BOVINE BLGS UPON THE ADDITION OF PNPP.

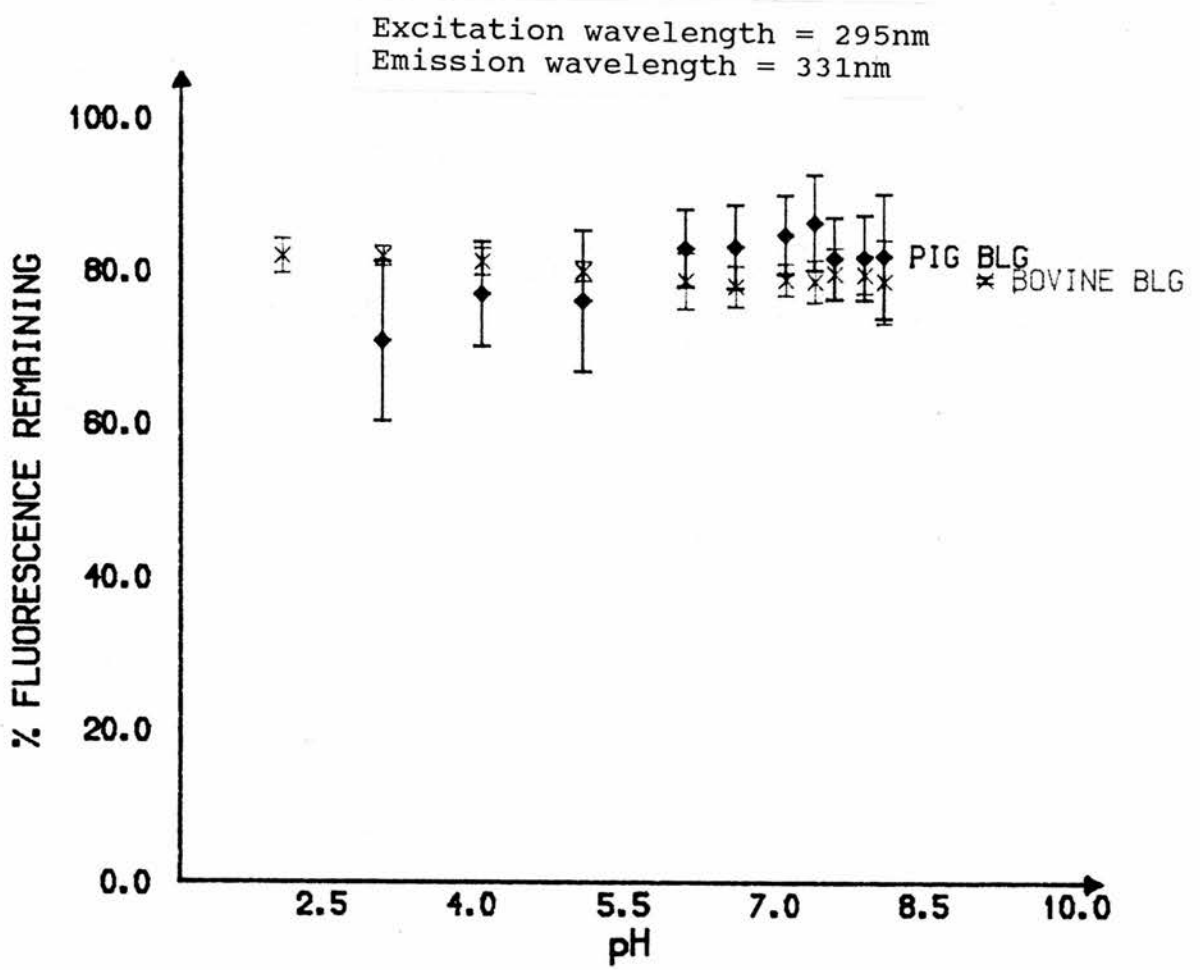
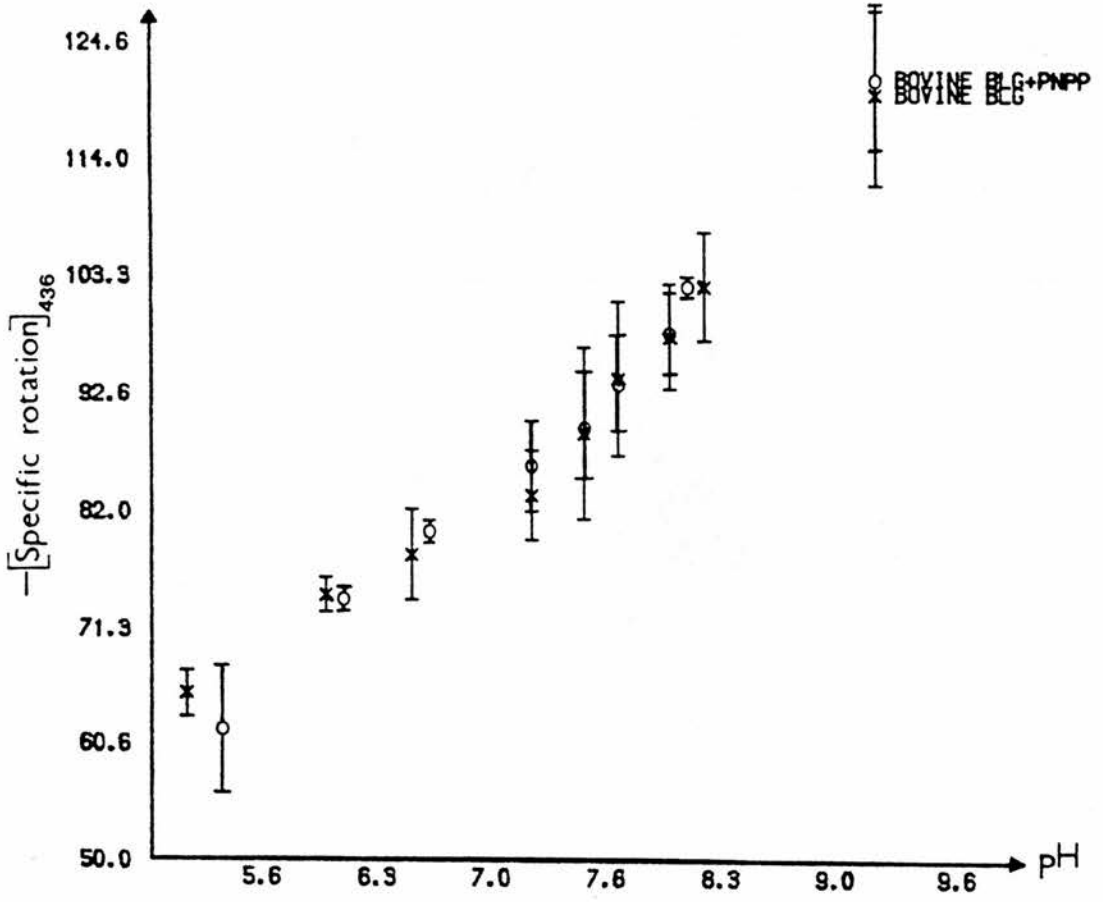


Figure 4:13

THE VARIATION IN SPECIFIC ROTATION OF BOVINE BLG UPON THE ADDITION OF PNPP.



4:8 BILIVERDIN BINDING TO BLG.

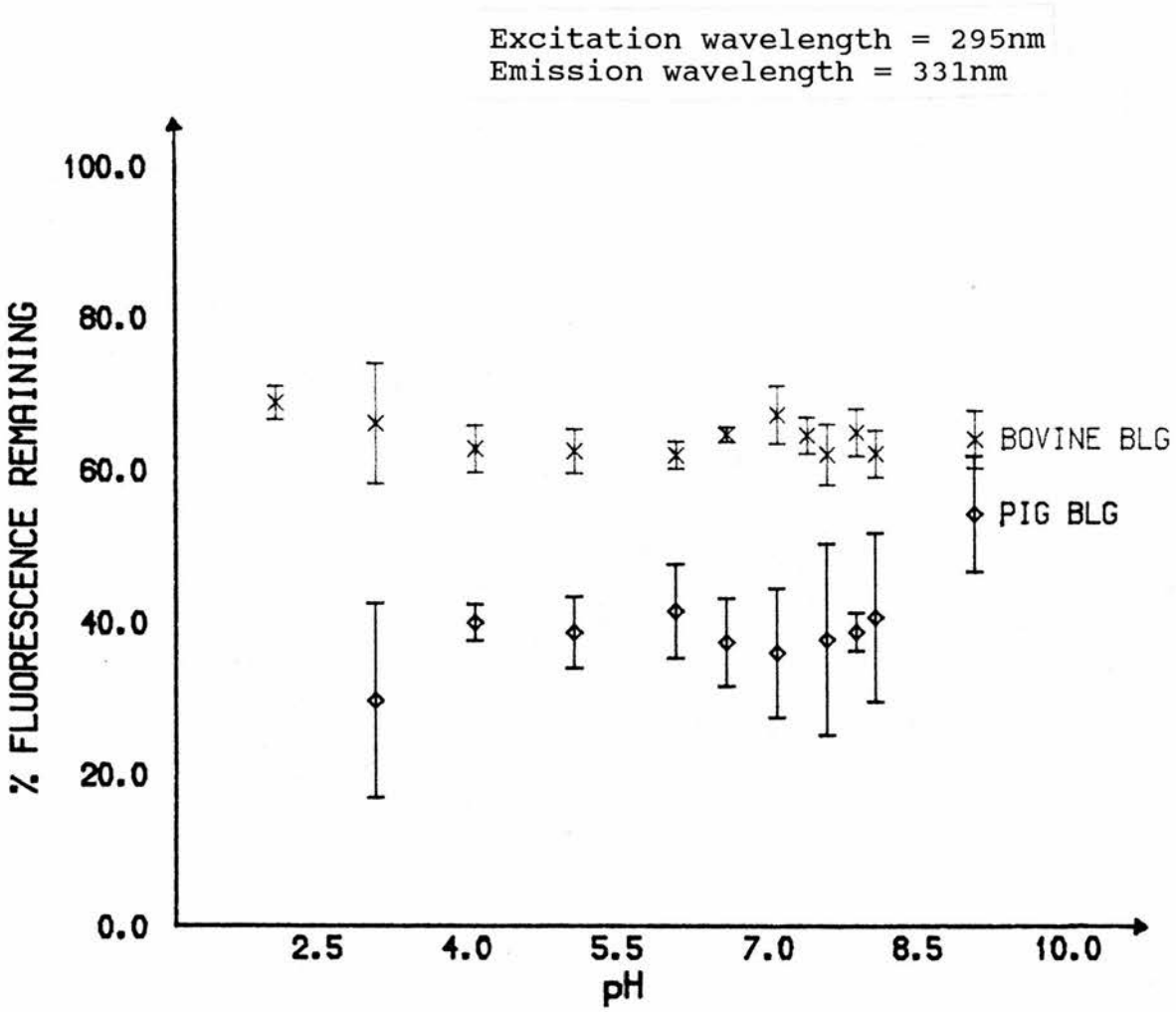
The structural homology between BLG, INCYN and BBP [222,223] suggested that BLG might bind the ligand of the other two proteins - biliverdin. Preliminary Trp fluorescence studies supported this view, by revealing decreases of about 36% and 60% in this fluorescence of bovine and pig BLGs respectively, upon the addition of this ligand {Figures 4:4 and 4:5}. Further studies were carried out to confirm the existence of a BLG+biliverdin complex, and provide information on the binding site.

The titration of bovine BLG (pH 7.3) with biliverdin was studied by Trp fluorescence {Figure 4:6}. The linearity of the Stern-Volmer plot implied that the fluorescence from only one type of Trp residue was being quenched, most probably that from the exposed Trp-61 near the mouth of the hydrophobic cavity. This appears reasonable as crystallography has shown that the binding site for biliverdin in both INCYN and BBP was around the mouth of the β-barrel [222,223]. The Kd value obtained from this fluorescence quenching was 9μM. The titration of pig BLG with biliverdin was not carried out, but would be of considerable interest as this protein does not contain Trp-61, and yet its fluorecence was quenched more than that of bovine BLG by this ligand. This seems to suggest that collisional quenching must occur via the region of monomeric pig BLG which is buried by dimerization in bovine BLG.

The nature of the biliverdin binding site was examined by Trp fluorescence and near UV CD spectroscopy, whilst the effect of this binding on the Tanford transition was monitored by polarimetry. The '% fluorescence remaining' upon the addition of biliverdin to bovine BLG decreased slightly from pH 2 to 6, but remained constant around 63% between pH 6 and 8 {Figure 4:14}. The latter constancy is probably due to the

Figure 4:14

THE QUENCHING OF TRP FLUORESCENCE FROM PIG AND BOVINE BLGS UPON THE ADDITION OF BILIVERDIN.



carboxyls in the surface region of the protein around Trp-61 (Glu-62, Asp-64 and Glu-65 in BLG-A) being fully ionized rather than the binding site being hydrophobic. The addition of biliverdin to pig BLG reduced the '% fluorescence remaining' to around 36% {Figure 4:14}, although this value appeared to increase across the pH range 6.5 to 8.5. This suggests that the association of biliverdin with pig BLG is different to that found with bovine BLG, probably due to both its monomeric state, and variations in the loops near the mouth of the β-barrel where biliverdin is believed to bind.

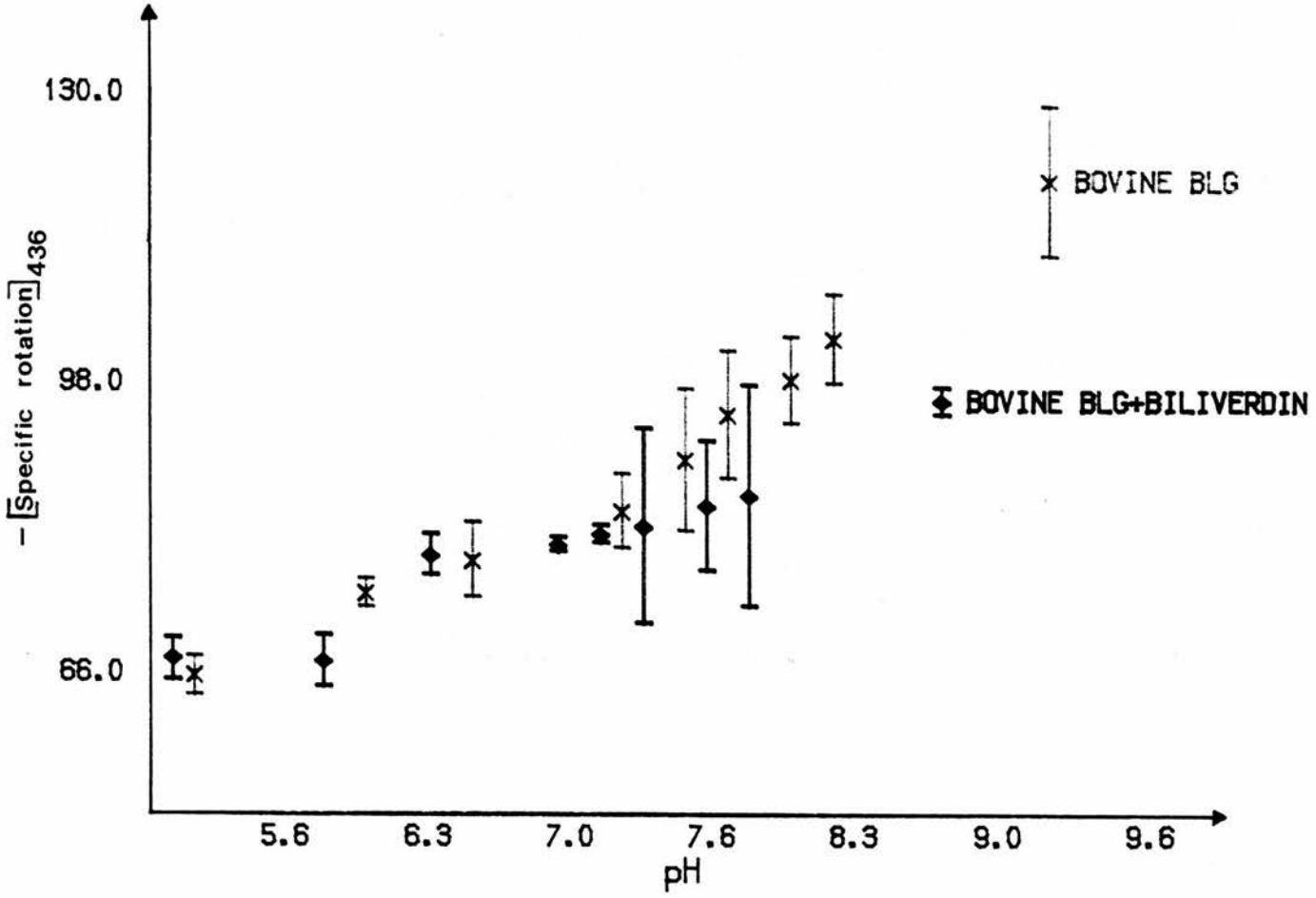
Near UV CD spectra were obtained for bovine BLG, in the presence and absence of biliverdin, at pH values 6 and 8 {Figure 4:8}. A saturated solution of biliverdin in phosphate buffer at pH 9 was used, and the true concentration of the ligand determined by UV absorption spectroscopy at 380nm. Despite this the samples never contained an excess of biliverdin, as the concentration of protein would have been too low to detect. At pH 6 the signals at 285nm (Tyr) and 293nm (Trp) were affected by the ligand, implicating these types of residues in the binding region. Trp-61 and Tyr-42 are the most likely candidates as they are close to one another, whereas Trp-19 and Tyr-20, Tyr-99 and Tyr-102 are all located towards the foot, rather than the mouth of the β-barrel. At pH 8 the ellipticities of the bands at 267nm (Phe), 285nm (Tyr) and 293nm (Trp) are almost doubled by the addition of biliverdin. This large change seemed odd, considering the low concentration of biliverdin present, and the constancy of the '% fluorescence remaining' values at pH 6 and 8. One possible explanation is that the use of EtOH (as the solvent for biliverdin) in the Trp fluorescence experiments prevented the pH becoming too alkaline, whilst the use of a phosphate buffer at pH 9 in the optical rotation experiments raised the pH too high in microenvironments, and caused denaturation of the protein

which is known to be less stable at higher pH.

A polarimetric study on bovine BLG+biliverdin (in phosphate buffer pH 9) revealed that the ligand had no effect on the specific rotation of the protein up to pH 7, which further supports the idea that binding is near Trp-61, and not Trp-19 which is close to the region involved in the Tanford transition. At higher pHs the magnitude of the specific rotation in the presence of biliverdin became less, and the error bars larger, which is consistent with the presence of some denaturation (Figure 4:15).

Figure 4:15

THE VARIATION IN SPECIFIC ROTATION OF BOVINE BLG UPON THE ADDITION OF BILIVERDIN.



CHAPTER 5

DISCUSSION AND CONCLUSIONS

5:1 INTRODUCTION.

When this project began in 1985 there was a wealth of information available on the solution properties of bovine BLG, but very few structural details with which to interpret these observations. A preliminary, medium resolution crystallographic structure had just been published [158,8], which identified the site of association between subunits in this dimeric protein, and revealed its structural homology to RBP.

The aim of our research was to improve the resolution of the structure available, and to use it, in conjunction with other medium-to-high resolution structures, to explain some of the properties of BLG. We wished to identify the environments of specific amino acids, locate the antigenic regions of this protein, and offer a molecular description for the conformational change (the Tanford transition) which occurs over the physiological pH range. Whilst this project was being undertaken the structural homology between BLG and RBP was extended to INCYN and BBP [222,223], and their sequence homology identified in many other proteins (Figures 1:15 and 1:16) [9]. This suggested that BLG was a member of a superfamily of transport proteins, which carried small, conjugated, hydrophobic/labile molecules; and was supported by the isolation of ruminant BLGs with 0.5 moles of FFA bound [240]. The probable importance of ligand binding for the function of BLG indicated that it would be worthwhile to investigate the nature of the different binding sites, and this was undertaken.

A summary of the work presented in this thesis is

now given, and discussed in terms of the current knowledge available. A possible function for BLG is then proposed, and the evidence for it is examined.

5:2 THE LATTICE Y STRUCTURE OF BLG.

5:2:1 DETERMINATION OF THE LATTICE Y STRUCTURE.

Orthorhombic crystals of bovine BLG lattice Y were obtained by salting-out the protein from a solution whose pH dropped from 7.8 to 7.3. A preliminary, medium resolution crystallographic structure was already available, the data having been collected on a conventional x-ray source, and phased using isomorphous derivatives [5,8]. Medium and high resolution data were collected on the synchrotron at Daresbury, and the former used to give an improved structure (R_F=48% and a figure of merit of 0.81). Refinement, using energy minimization and MD, reduced the R_F to 34.6%, whilst a Ramachandran plot revealed that there was still considerable room for improvement in the geometries of the residues {Figure 2:33}. Thus this structure is currently being manually model-built with geometrical contraints, prior to restrained LS refinement and the subsequent inclusion of our higher resolution data.

The lattice Y structure of bovine BLG consists of nine anti-parallel β-strands, of which eight form two β-sheets that wrap around one another to enclose a large hydrophobic cavity (about 360A^3). The ninth is involved in the association of subunits to give a dimer, and there is also a three-turn α-helix near the C-terminus of this protein {Figures 1:13 and 2:36}. The current resolution of this structure is adequate to enable the positions and environment of certain residues to be examined.

This structure has confirmed the presence of two

disulphide bridges (Cys-66 -> Cys-160 and Cys-106 -> Cys-119), and a free thiol at Cys-121. This is in complete agreement with both the 2A lattice X (pH 6.5) and 2.5A lattice Z (pH 7.5) structures, but does not exclude the possibility that disulphide interchange could occur in solution, producing equal amounts of BLG containing the disulphide bridges Cys-106 -> Cys-119 and Cys-106 -> Cys-121 [146,147]. Preferential crystallization of the first form could have occurred, but molecules in the other conformation would have had to convert to the first form, in order to explain how very much more than 50% of the protein present crystallized in the observed form. However this conversion seems unlikely as a large shift in the position of Cys-121, together with the breaking and reformation of about 10 H-bonds would be required {Chapter 2:8:5}.

The environments of the Trp and Tyr residues had been investigated previously by solution studies [150], but it was now possible to use our crystallographic structure to confirm these observations and identify the specific residues involved. Trp-19 is located at the foot of the hydrophobic β-barrel on the inner side (as observed in RBP), whilst Trp-61 was in an ill-defined flexible loop near the mouth of the cavity {Figure 2:43}. Of the Tyr residues Tyr-42 is fully accessible and believed to be part of one of the antigenic sites of BLG. Tyr-20 is also fully exposed, and, together with the partially buried Tyr-99, is part of the region of the protein conserved in BLGs and members of the BLG superfamily, which may be important for binding the protein to a receptor {Figures 2:37} [228]. Tyr-102 is fully buried, associated with the changes in Phe-136 and Cys-121 over the Tanford transition {Figure 3:9}, and may be part of the external hydrophobic binding channel.

5:2:2 DIMER AND OCTAMER ASSOCIATION SITES.

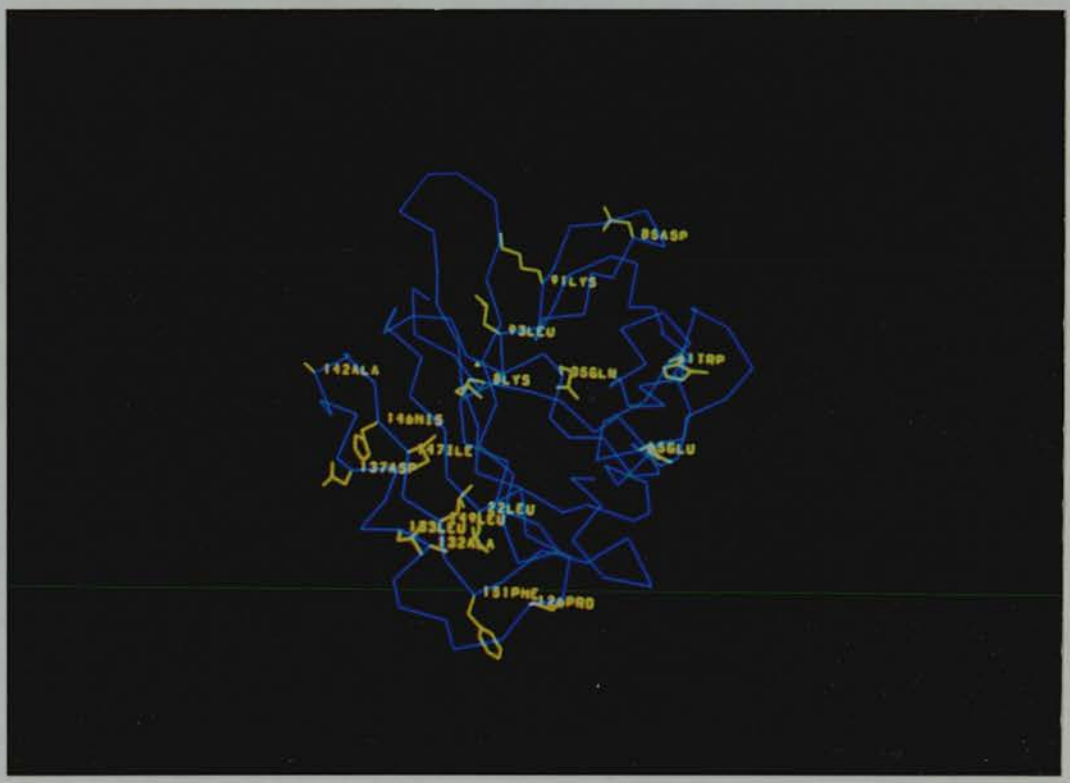
The site of association between subunits in dimeric bovine BLG is reported to involve hydrophobic interactions between Ile-29, Ile-147, and the stacking of symmetry-related imidazoles of His-146 (Figure 2:39) [8]. However, although these residues are conserved in all known dimeric BLGs they are not absent from all monomeric BLGs (Figure 1:11) - dolphin BLG contains all three, but is monomeric [44]. The involvement of Ile-29 has been confirmed, as a mutant ovine BLG containing Ser-29 has been prduced by site-directed mutagenesis, and is known to be monomeric [93].

There are only 20 amino acids which are conserved in all dimeric BLGs, but substituted in all monomeric proteins and the positions of many of these are illustrated in Figure 5:1. They are located in three main regions of the structure: the N-terminal region (residues 2, 8, and 9), the loops around the mouth of the hydrophobic cavity (residues 35, 61, 65, 85, 91, 93, 101, and 110), and in the α-helix and random-coil at the C-terminus of the protein (residues 132, 133, 137, 151, and 152). The significance of these regions is unclear - although the latter residues are located near the site of association. The closeness of His-146 and Ile-147 to the free thiol (Cys-121 -> His-146 is 12.7A whilst Cys-121 -> Ile-147 is 11.7A) may explain why thiol-specific reagents enhance the dissociation of dimeric BLG.

Octamerization has only been observed in bovine BLG variant A near pH 4.5 and $0^{o}C$ [94,98]. Primary sequence homology between BLG-A and B implied that Asp-64 or Val-118 could be involved, but the latter is unlikely as it is buried within the structure. A comparison between variants A and Dr revealed that the substitution of Asp-28 by Asn prevented octamerization, but allowed glycosylation of the latter protein. The crystal

Figure 5:1

PHOTOGRAPH TO SHOW THE POSITIONS OF THOSE RESIDUES CONSERVED IN ALL DIMERIC, BUT SUBSTITUTED IN ALL MONOMERIC, BLGS.



structure has revealed that the distance between Asp-28 and Asp-64 is 23.5A, and both are located on flexible loops at opposite sides of the mouth of the hydrophobic cavity. Hence Asn-28 in BLG-Dr is accessible for glycosylation, although a carbohydrate moiety would be likely to obstruct the entrance to the β-barrel. As its side chain points in towards the entrance to the cavity, whilst Ile-29 is directed outwards into the interface between subunits of the dimer, this could explain why there appears to be no conflict between glycosylation and dimerization in BLG-Dr.

5:3 THE TANFORD TRANSITION OF BOVINE BLG.

The Tanford transition is defined as the reversible change in optical rotation observed when the pH of bovine BLG varies from 6.5 to 7.8 [6]. The slight decrease in the sedimentation coefficient observed across the physiological pH, is mainly due to a change in the conformation of the protein causing an expansion of its volume, and not increased dissociation of this dimeric protein {Chapter 3:1}.

The nature of this transition was studied in solution, and then a comparison was made between the crystallographic structures of lattice X at pH 6.5 [141], and lattice Y at pH 7.8, so as to offer a molecular explanation for this conformational change. Solution studies have revealed that the Tanford transition is a rapid conformational change, taking less than 10 seconds [312]. This quick response of BLG to a variation in pH suggests that the transition could occur <u>in vivo</u> as the pH increases down the GI tract prior to its binding to specific receptors which have been detected in the lower third of calf ileum [8]. Upon raising the pH from 6 to 8 there appears to be virtually no change in the secondary

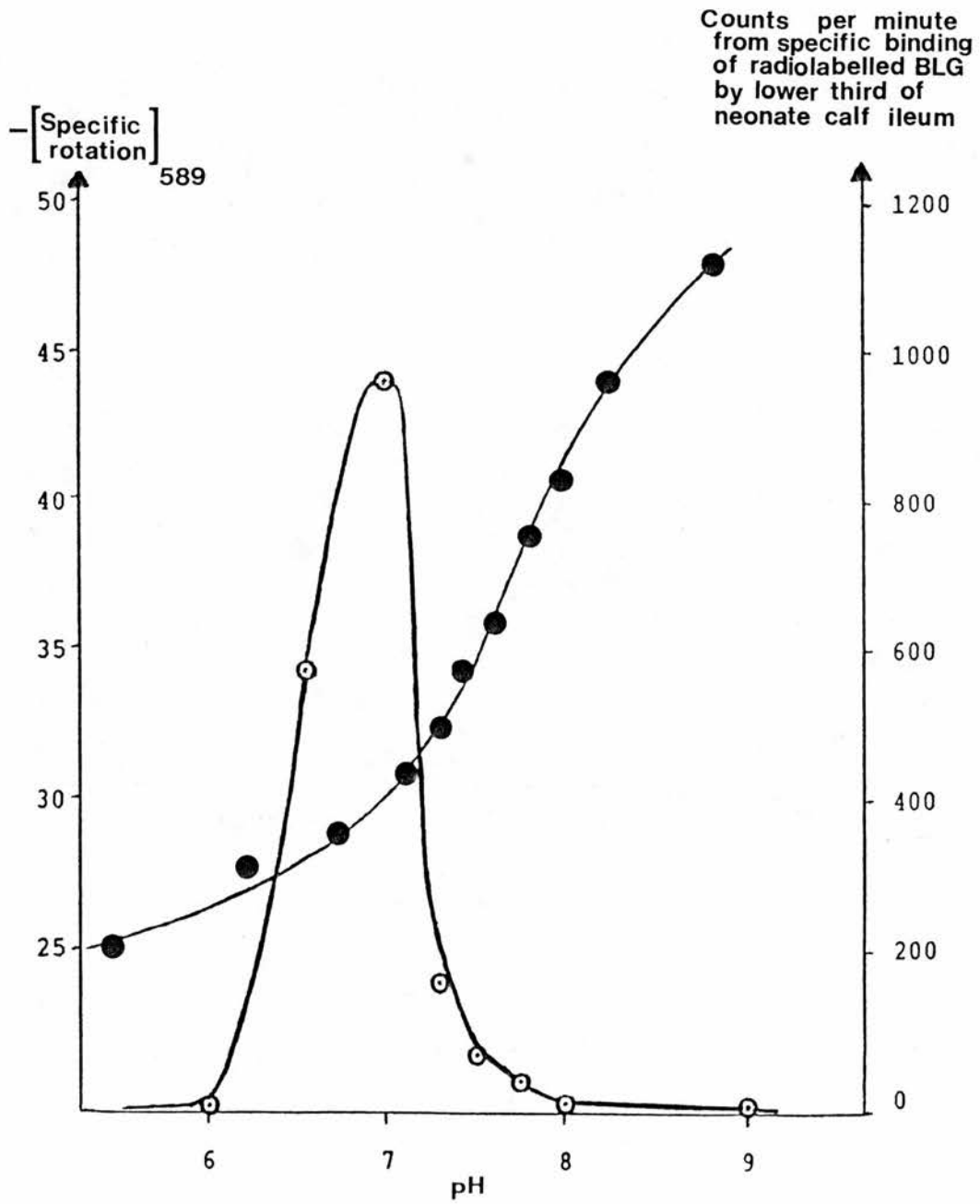
structure of the protein {Figure 3:3} - the β-barrel remains intact, implying that any ligand being carried within the hydrophobic cavity is still well protected and unlikely to be released by this transition; whilst the slight decrease in α-helical content at the higher pH could be explained by the reorientation of Phe-136 between lattices X and Y {Figure 3:9}. The movement of this Phe appears to correlate with changes in the positions of Tyr-102 and Cys-121. The latter is shifted 3.4A and becomes more exposed upon raising the pH, which explains the increased reactivity of BLG to thiol-specific reagents at higher pH [138,139]. All three residues are located near one another (Cys-121 -> Tyr-102 is about 2A, and Cys-121 -> Phe-136 is around 5A), suggesting that the Tanford transition may be a localized conformational change, and possibly necessary for binding to a receptor.

Near UV CD spectroscopy has also implicated Trp residues in this transition {Figure 3:4}. Trp-19 is probably involved as it is located in the same region of the protein (Trp-19 -> Cys-121 is 8A), and changes orientation between lattices X and Y, becoming more buried in the latter {Figure 3:8}. The inclusion of Trp-61 is unclear, as its position is not defined in lattice X. As Trp-19 is thought to be important for the binding of molecules to BLGs [153,164] this implies a link between the Tanford transition and ligand binding. The addition of retinol to bovine BLG did shift the midpoint of the Tanford transition to a lower pH (7.2) {Figure 3:15}, which is nearer that at which maximal uptake of retinol from the BLG+retinol complex by specific receptors is observed {Figure 5:2}. However no change was detected when retinol was added to pig BLG {Figure 3:16} - suggesting that the shift observed with bovine BLG is not of functional importance.

The variation in the orientations of Phe-136, Tyr-

Figure 5:2

RELATIONSHIP BETWEEN THE VARIATION WITH pH OF THE UPTAKE OF RETINOL BY CALF ILEUM AND THE TANFORD TRANSITION.



KEY

● = ORD measurements

O = scintillation counts

102 and Trp-19 over the Tanford transition of bovine BLG are sufficient to account for the change in optical rotation observed. In pig BLG Tyr-102 is replaced by His, and this may explain the smaller change in optical rotation. The fact that a conformational change occurs in both pig and bovine BLG seems to indicate that the Tanford transition may be a conserved feature of BLGs, and thus possibly important for their function.

The Tanford transition is known to be linked with the exposure and ionization of a single anomalous carboxyl per subunit (pK 7.4). It is not clear whether the increase in pH from 6 to 8 causes the exposure and ionization of the carboxyl resulting in the reorientation of certain residues (Phe-136, Tyr-102 and Cys-121), or vice versa. A comparison of the lattice X and Y structures revealed three carboxyls, Asp-33, Glu-74 and Asp-96, which were located in the interior of the protein and involved in H-bonding at pH 6.5, but were in different, more exposed positions around pH 7.8 (Figures 3:10 and 3:11). Any one of the three could show anomalous behaviour, although only Asp-96 lies near the other residues associated with the Tanford transition. As site-directed mutagenesis of the ovine BLG gene can now be carried out, the most promising way to unambiguously solve the identity of the anomalous carboxyl seems to be the replacement of each of these candidates in turn, and then polarimetric studies on the mutated proteins.

5:4 BINDING SITES IN BLG.

Bovine BLG is known to bind a wide variety of compounds (Chapter 1:7 and Figure 1:12). The binding site for the large protein molecules which associate with BLG is believed to be different to those for smaller organic ligands.

Four discontinuous antigenic regions of bovine BLG were proposed from solution studies, and some of the important residues in these regions identified {Figure 2:42} [209]. Two of these regions, 62-107 and 146-162, which are spatially adjacent and linked by the disulphide bridge Cys-66 -> Cys-160, may be part of a single site. The region 124-140 is currently of interest, as sequence homology has been detected between these residues and sections of β_2-microglobulin [33] and lactoferrin [22b] {Figure 5:3}. Antibodies to bovine BLG are known to cross-react with each of these proteins, but it seems likely that different residues within this region are involved, as only a single Thr residue is conserved between all three. The failure of normal titres of antibodies against bovine BLG to react with monomeric BLGs, eg. pig [83b], is surprising considering the high sequence homology between these BLGs compared to that between bovine BLG, β_2-microglobulin and lactoferrin.

Cytochrome-c can form a 1:1 complex with bovine BLG [183], but the nature and site of association has not been previously identified. The CYT-C structure is known to contain a ring of basic Lys residues [190], whilst the positions of some acidic residues in BLG were also discovered to form a ring centred around Cys-121, with a 'radius' varying from 6.9A for Cys-121 -> Asp-129 to 11.9A for Cys-121 -> Glu-137 {Figure 5:4}. This suggests that the interaction between these two proteins may be electrostatic, with Cys-121 near the middle of the interface. The conformational change required, for the free thiol in bovine BLG to be able to reduce the iron in CYT-C at pH > 7.5, is likely to be the Tanford transition which is known to enhance the accessibility of the thiol.

The binding of 2,6-MANS to bovine BLG-B indicated the presence of two anionic binding sites, which were hydrophobic [162]. These sites were unequal, the stronger having an association constant(Ka) of $3.4 \times 10^5 M^{-1}$.

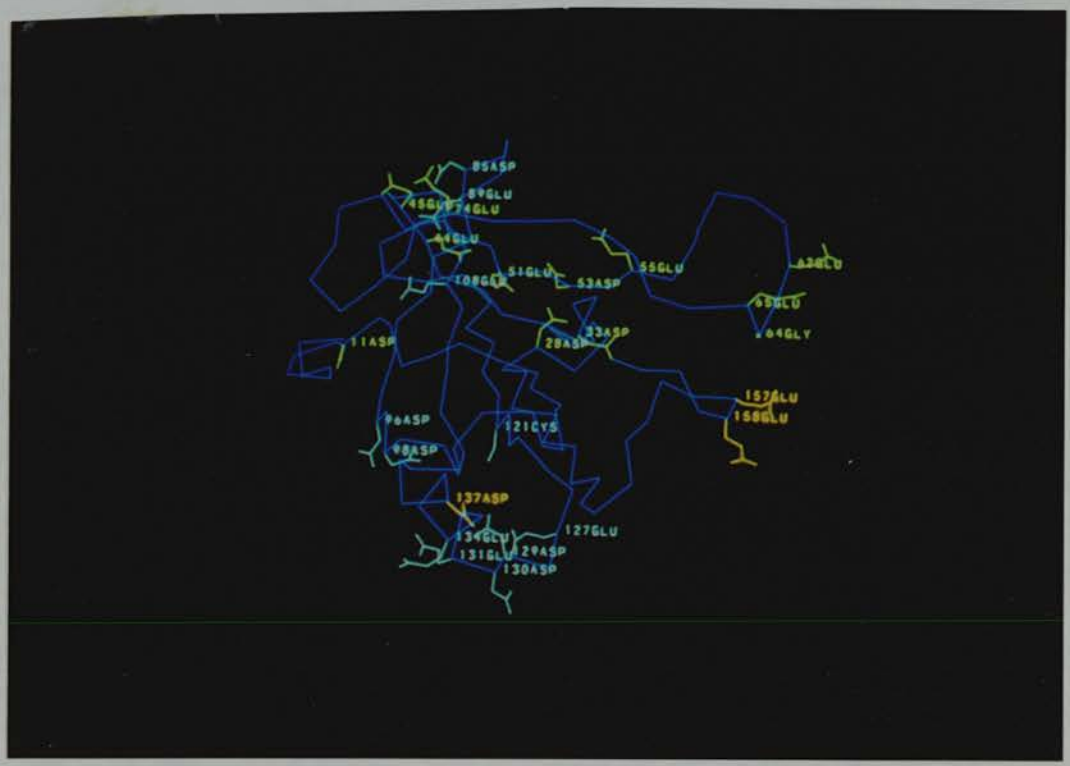
Figure 5:3

HOMOLOGOUS REGIONS BETWEEN BLG, β_2-MICROGLOBULIN AND LACTOFERRIN.

BOVINE BLG (124-140)	R	T	P	-	-	E	V	D	D	E	-	A	L	E	K	F	D	K	A	L
β_2-MICROGLOBULIN (71-86)	-	T	P	-	T	E	K	D	-	E	Y	A	C	-	R	V	N	H	V	T
LACTOFERRIN (121-135)	R	T	A	G	W	N	V	P	I	G	-	T	L	R	P	F				
PIG BLG (124-140)	R	T	L	-	-	E	V	D	D	Q	-	I	R	E	K	F	E	D	A	C

Figure 5:4

LOCATION OF THE ACIDIC RESIDUES AND CYS-121 IN BOVINE BLG.

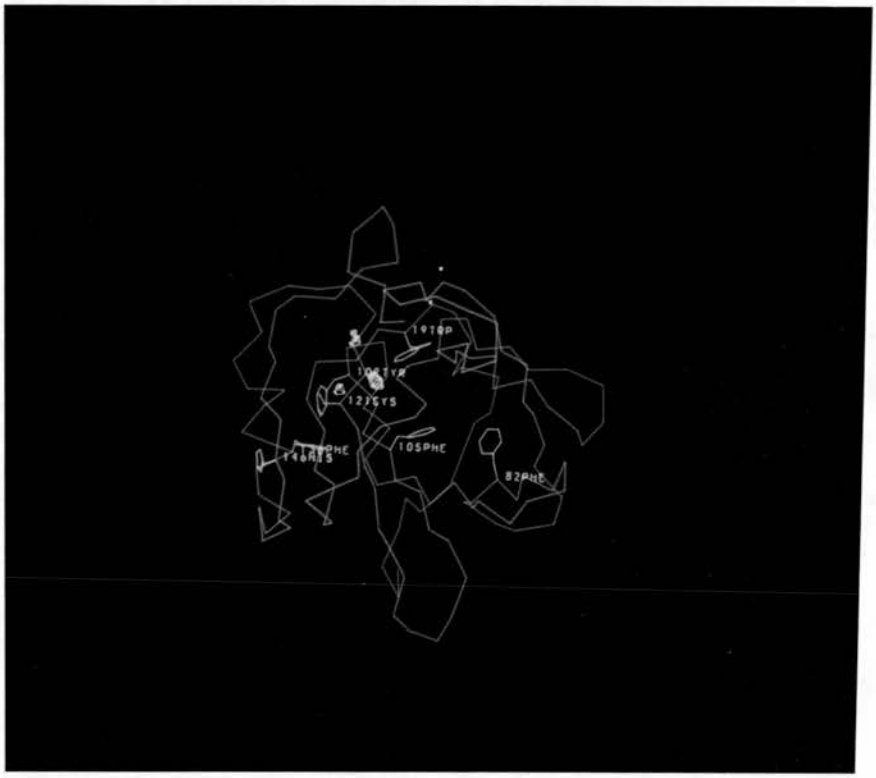


Analogous to this is the binding of FFAs to BLG - one molecule of palmitic acid can bind at a primary site with a Ka of $10^5 M^{-1}$, whilst up to six more can bind at a secondary site with a Ka of around $10^3 M^{-1}$ [176]. A more detailed analysis of the binding of one molecule of palmitic acid to BLG revealed that the Ka increased across the pH range 6.5 to 8.7, that there was an increase of only 8% in the Trp fluorescence of this protein upon binding, and that there was a cationic residue at the mouth of the binding site. The two probable binding sites in BLG have been identified as the hydrophobic cavity with Trp-19 at its foot and Lys-70 near its mouth, and the external hydrophobic channel involving the α-helix and Cys-121 with Lys-141 near its opening. This, together with our knowledge, that the latter site is linked with the Tanford transition {Chapter 3:3}, and is not far from Trp-19 (about 9A), seems to imply that the primary FFA binding site is the external channel.

As FFA, SDS, iodobutane and alkanes have all been proposed to bind at the same site [169], this indicates that the 'alkane binding site' is probably equivalent to the external channel. Further support for this comes from a crystallographic study carried out on bovine BLG lattice Y crystals soaked in iodobutane. The data was collected, and processed by M.Z.Papiz, but scaled to the resultant native BLG dataset, up to 2.8A resolution, by ourselves - a R_{scale} of 27.1% was obtained for the 3940 common reflections used. An electron density difference map, with coefficients $(2F_{BLGI} - F_{BLG})$ was generated, and revealed that extra electron density was located near Asp-98, Tyr-102 and Cys-121 {Figure 5:5}. The binding of iodobutane in the external channel seems to confirm that this is the alkane binding site.

Figure 5:5

PHOTOGRAPH TO ILLUSTRATE THE STRONGEST AREAS OF ELECTRON DENSITY IN THE 2.8A DIFFERENCE MAP OF (2(BLG+IODOBUTANE) - BLG).



5:4:1 RETINOL BINDING SITES.

The binding of retinol to BLG has been studied [153,160,161], but the site of binding was not identified. However there are several observations to suggest that retinol does not bind in the β-barrel of BLG in a manner analogous to RBP : rotational times suggest that the retinol must be in a flexible part of the protein [153], alcohol dehydrogenase degrades retinol bound to BLG, but not retinol bound to RBP [161], and the position of the α-helix in BLG is similar to that found in RBP after the loss of retinol [218]. We propose that the retinol binds in the external hydrophobic channel near residues Phe-136, Tyr-102 and Cys-121, and not far from Trp-19 (about 9A), in a manner analogous to palmitic acid. However why neither ligand can bind in the cavity is unclear, although as the protein was not treated with hexane randomly orientated hydrophobic molecules could still occupy the barrel. Support for binding in the channel has been obtained from crystallography, polarimetry, CD and fluorescence studies.

X-ray crystallography of the BLG lattice Z+retinol complex confirmed the ligand was located in the external channel [59], whilst our comparison of the structures of lattice X (pH 6.5) and Y (pH 7.8) showed that those residues which change orientation over the Tanford transition (Phe-136, Tyr-102 and Cys-121) were positioned in the same region of the protein {Chapter 3:3}. Polarimetry confirmed that this transition in bovine BLG was affected by retinol binding {Figure 3:15}. However in pig BLG the effect of retinol binding on the transition could not be clearly detected as the transition itself was considerably smaller (30% of that observed in bovine BLG), due to the substitution of Tyr-102 by His {Figure 3:16}. CD studies, using solutions only slightly greater than 1:1 retinol:subunit of BLG, revealed that for bovine

BLG at pH 6, Phe but not Tyr residues were affected by the addition of retinol - this correlates well with the crystallographic structure which shows that Tyr-102 is positioned deeper in the channel than Phe-136. At pH 8 pig BLG, upon the addition of retinol, indicated that both Tyr and Trp residues were involved, suggesting that the environment of the channel was altered by the Tanford transition. This could enable closer contact between retinol, Tyr-99 and Trp-19, the latter being known to become less accessible to solvent at the higher pH end of the Tanford transition {Chapter 3:3}.

In a manner analogous to FFA binding to BLG, the first molecule of retinol which binds in the external channel may open up the β-barrel of BLG for the binding of one or more additional retinol molecules in the cavity. This closer approach of retinol to Trp-19 could explain the enhanced quenching of the Trp fluorescence of bovine BLG {Figure 4:6}. No enhanced quenching was observed when excess retinol was added to pig BLG - one explanation could be that the β-barrel of pig BLG differs enough to prevent retinol getting to its foot, and the substitution of Leu-104 in bovine BLG by Phe in pig BLG could be significant.

Fluorescence studies were carried out on both pig and bovine BLG, upon the addition of retinol, and showed a decrease in fluorescence from both BLGs, due to the bound retinol in the channel being able to quench Trp-19 fluorescence. Upon raising the pH across the Tanford transition the quenching was enhanced, which is consistent with increased accessibility of the retinol to the base of the channel and stronger binding (stronger binding was also observed across the pH range 6.5 to 8.7 upon the binding of palmitic acid to bovine BLG [176]). From an examination of the lattice X and Y structures it seems likely that the movement of Phe-136 and Tyr-102 could enable the β-ionone ring to become positioned near

both rings at the higher pH (Figure 3:9). The absence of Tyr-102 could explain the smaller increase in the quenching over the Tanford transition, and the larger dissociation of the pig BLG+retinol complex (Kd is about 10μM).

Retinol fluorescence also seemed to be related to the interaction of Tyr-102 with the ligand - a large enhancement, which increased across the Tanford transition, was observed when bovine BLG was added to retinol. Negligible change was seen when pig BLG was used. This suggests that Tyr-102 plays a major part in the Tanford transition and retinol binding in the external channel. Site-directed mutagenesis of this residue and subsequent analysis of the mutated protein by the above techniques should help clarify many of the observations and comments made.

5:4:2 PNPP BINDING SITE.

Our intention was to study the binding of PNPP to bovine BLG by both crystallography and solution studies. Crystallography was carried out on the yellowish crystals obtained by soaking BLG lattice Y crystals in PNPP, and revealed that the p-nitrophenyl ligand was located within, and near the foot of, the β-barrel (Figure 4:10). However it seemed unlikely that the phosphate anion could favourably exist within the hydrophobic cavity - it was more probable that the ligand within the barrel was p-nitrophenol, having been formed from the hydrolysis of PNPP during the soaking. This is reasonable as p-nitrophenol is also known to form a complex with BLG (Kd of 52μM), and would account for the yellow colour of the crystals.

Solution studies support the binding site proposed from crystallography. Polarimetric studies showed that the binding of PNPP to BLG did not affect the Tanford

transition (Figure 4:13), whilst Trp fluorescence studies showed that the decrease upon PNPP addition was independent of pH, both of which seem to exclude binding in the external channel. This correlates with the crystallographic observation that p-nitrophenol binds in the hydrophobic cavity.

5:4:3 BILIVERDIN BINDING SITE.

The binding of biliverdin to BLG was studied, as this ligand is known to bind to the structurally homologous proteins INCYN [222] and BBP [223]. Solution studies seemed to confirm that biliverdin could bind to bovine BLG in a similar position to that at which it was located in the bilin binding proteins by crystallography ie. at the mouth of the hydrophobic cavity. This appeared reasonable as biliverdin is rather large to fit within the hydrophobic cavity, and unlikely to unfold into a long chain for binding in the channel as this would involve the conversion of some cis bonds to trans ones.

The linear Stern-Volmer plot suggested that the fluorescence from only one Trp, Trp-61, was quenched, and indicated that the Kd value was 10μM. Trp fluorescence studies showed that the quenching was independent of pH, indicative of a non-ionizable binding site - this is consistent with that around Trp-61 where all the carboxyls would be ionized. CD spectra obtained at pH 6 implicated Tyr and Trp residues in the interaction, and from our BLG structure the most likely candidates are Tyr-42 and Trp-61.

5:5 COMPARISON OF PIG AND BOVINE BLGS.

A comparison was made between the primary sequences, the available structural information and some of the

properties of monomeric pig, and dimeric bovine BLGs, in order to elucidate those conserved features of BLG which could be of functional importance.

5:5:1 PRIMARY SEQUENCE ALIGNMENT.

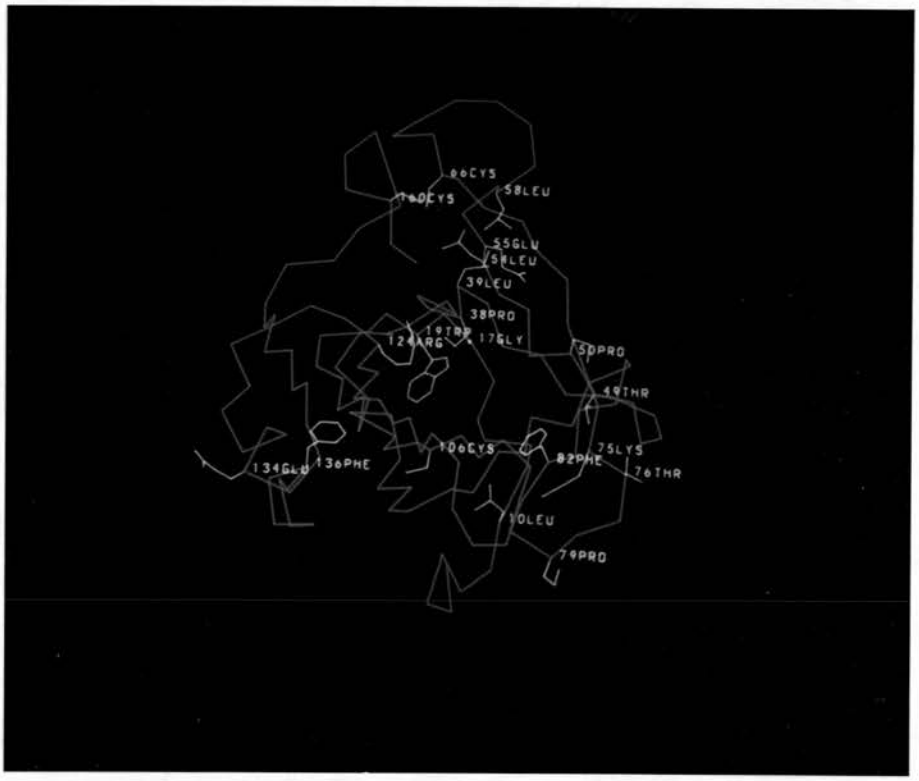
The primary sequences of several monomeric and dimeric BLGs, including pig and bovine, have been aligned previously, and the many common residues noted [85]. Our inclusion of additional BLGs (Figures 1:10 and 1:11) reduced the number of conserved amino acids to around 20, and their positions are illustrated in Figure 5:6. There was a wide spatial distribution of these residues across the structure; from Gly-17, Trp-19 and Phe-82 in the β-barrel to Glu-134 and Phe-136 in the α-helix, and from the residues of the C-terminal disulphide, Cys-66 and Cys-160, to several scattered in different loop regions. Thus these conserved residues failed to provide any strong evidence as to the functionally important regions of BLG. It was of interest to note that the region TDY--Y (97-102) was not conserved within all BLGs, despite being one of the conserved sequence regions in the superfamily of proteins. However, with the exception of horse BLG-II, the nature of the residues in this region was maintained in the BLGs, implying that a sequence comparison may be more informative if both the identity and nature of the amino acids is considered.

5:5:2 STRUCTURAL COMPARISON.

The most usual way of obtaining a structural comparison between two proteins is by determining both their structures by crystallography. However suitable crystals of pig BLG could not be grown under the same crystallization conditions as bovine BLG. Microcrystals of pig BLG have been reported [82], but reproducing those

Figure 5:6

PHOTOGRAPH TO SHOW THE POSITIONS OF THOSE RESIDUES CONSERVED IN ALL DIMERIC AND MONOMERIC BLGS.



conditions at lower temperatures failed to slow the crystallization down enough to yield suitable crystals for diffraction.

Three other methods are available for gaining an insight into the structure of pig BLG - secondary structure prediction, CD, and modelling based on the known bovine BLG structure. The prediction of the secondary structure of both pig and bovine BLGs was kindly carried out by L.Sawyer, and the results, which are a combination of eight separate prediction methods are given in Figure 5:7. They reveal that the methods used are poor as they overpredict the α-helical content of bovine BLG. However a comparison between the pig and bovine predictions could provide a clue to any differences between the structures. These suggest that pig BLG may contain some α-helix near the N-terminus, and that the α-helical region, 130-140, found in bovine BLG may be extended to 125-140 in pig. This correlates with a slight reduction in sheet predicted for pig compared to bovine over the region 125-130. However overall the β-sheet predictions of both proteins are similar, but low compared to the amount known to exist in the bovine BLG structure (about 50%).

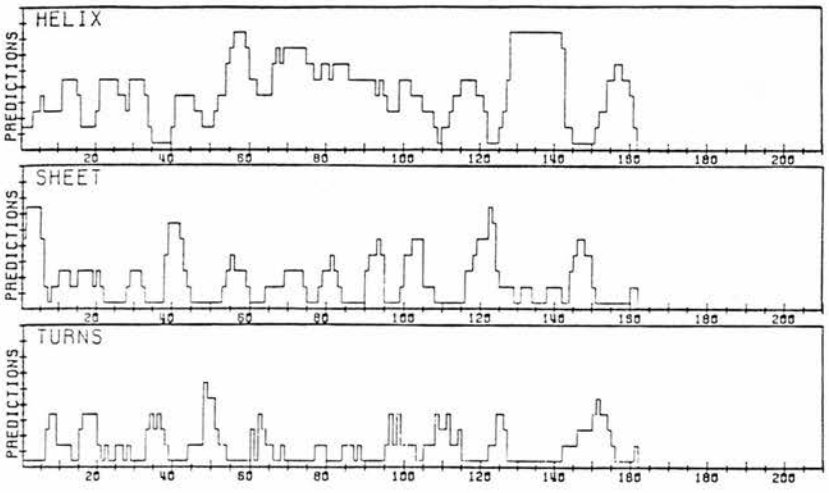
The far UV CD spectra of bovine and pig BLG at pH 6 showed very similar profiles, but were distinct (Figures 3:3 and 3:13). The negative band of pig BLG occurred at 208nm, whilst that of bovine BLG was broader with its minimium nearer 215nm, implying that pig BLG did contain a little more α-helix than the bovine protein.

From these predictions and observations it seems reasonable to propose that the β-barrel is maintained in monomeric and dimeric BLGs, and may be important for the function of the protein. By assuming structural homology between pig and bovine BLG computer graphics were then used to model a structure for the pig protein [212]. The main observation made was that the deletion of Phe-151

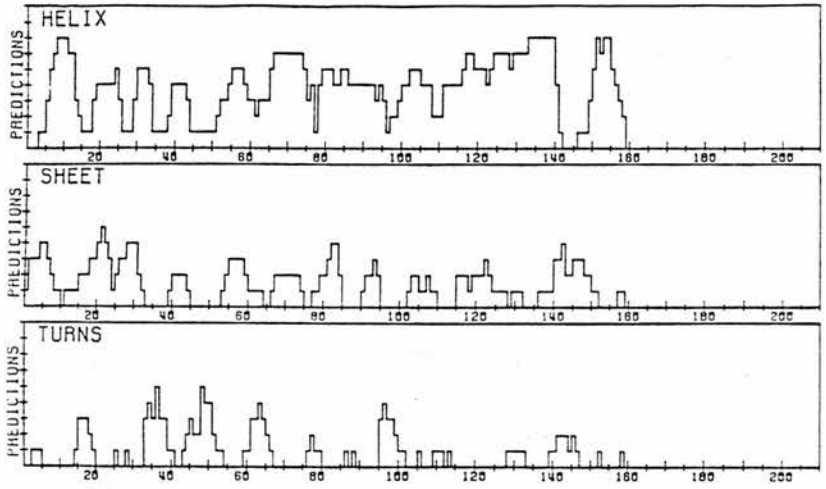
Figure 5:7

SECONDARY STRUCTURE PREDICTIONS FOR BOVINE AND PIG BLGS.

For Bovine BLG



For Pig BLG



and Asn-152 in pig BLG extended the adjacent residues bringing strand I (145-150) nearer to the C-terminus - Cys-66 -> Cys-160 was reduced from 2.0A to 1.4A. In bovine BLG strand I is parallel to the surface of the molecule and is known to be involved in dimerization by making anti-parallel β-strand interactions with the dyad-related strand; whereas in pig BLG strand I is perpendicular to the molecular surface, and may explain why this protein is monomeric.

5:5:3 DIFFERENCES IN SOLUTION PROPERTIES.

Various solution studies have revealed clues about the slight differences between the structures of pig and bovine BLGs. The near UV CD spectra of these proteins differed considerably {Figures 3:4 and 3:14}, but these differences could not be assigned to changes in the environments of specific residues as the numbers of Phe, Tyr and Trp residues also varied. Pig BLG has 4 Phe, 2 Tyr and 1 Trp compared to bovine which has 4, 4, and 2 respectively.

The secondary structure elements of these two proteins also differ slightly. The β-barrel is believed to be modified in pig BLG, such that ligands cannot approach as near to Trp-19 at the foot of the calyx. This was deduced from the lack of enhanced quenching upon the addition of excess of retinol to pig BLG, and the smaller decrease in Trp fluorescence observed upon the addition of PNPP to pig compared to bovine BLG. The α-helix has been predicted to be longer in pig BLG and this could alter the optimum length of a hydrocarbon chain which might bind in the external channel. As yet the binding of a homologous series to BLG has not been studied. The substitution of Tyr-102 in bovine BLG by His in pig seems to affect both the Tanford transition and the binding of ligands in the hydrophobic channel, but this residue is a

feature of pig BLG, rather than monomeric BLGs in general.

The differing decreases in the Trp fluorescence of pig and bovine BLGs upon the additions of L-dopa, vitamin K1 [243] and vitamin K2 seem to confirm slight differences in these proteins, which was expected from the low primary sequence homology. However until their binding sites are identified few more details can be deduced from these results.

5:6 THE BLG SUPERFAMILY OF PROTEINS.

The structural homology between RBP [215], INCYN [222], BBP [223] and BLG [8] suggested that these proteins might be related. The conserved regions of their amino acid sequences were then looked for in other proteins, and a superfamily obtained. Although the common regions (9-23, and 94-104) and the residues 124, 66, and 160, make up only about 18% of the protein sequence, they are all located in a small region of BLG {Figure 2:37}. It has been suggested that this area of the protein may be involved in the binding of BLG to its receptor [228]. The 16 members of this superfamily are listed, and their properties summarized in Figure 1:16. They are all 18-20kDa secretory proteins, and many are known to bind small hydrophobic/labile ligands and transport them to a receptor. The gene structures of five of the members have been determined [275], and they show homology to one another, having their intron/exon boundaries in the same place which is an indication of an evolutionary relationship.

However, further examination of these proteins is beginning to suggest that some subdivision of this superfamily may be constructive. The phylogenetic trees from the differences between the amino acid sequence of

these proteins were obtained [245], and have suggested that they may be subdivided into five groups:

(A) CRCYN, INCYN, BBP, APOD

(B) PURP, RBP

(C) MUP, OBP, BLG, PP14

(D) BG, HC, ESP, C8G

(E) A1AG

Subsequent studies have only partially supported these groups. It has recently been shown that APOD can form a 1:1 complex with bilirubin [244], suggesting that it may be the mammalian form of a bilin binding protein, and hence justifying its close relationship with INCYN and BBP. In addition these three proteins, together with CRYCN, are the only four in the superfamily which form a disulphide bridge between Cys-34 and Cys-160 rather than Cys-66 and Cys-160. It has been suggested that APOD may have a similar function to A1AG, the former being a locally produced acute phase protein, whilst the latter is an acute phase reactant. However there is no other strong evidence to justify the inclusion of A1AG in the first group.

The second group contains only RBP and PURP, both of which bind and transport retinol <u>in vivo</u>, and show higher than average sequence homology for any two members of this superfamily. The exclusion of BLG from this group suggests that its function is not as a retinol transport protein. This is supported by the observation that retinol can bind to bovine, pig, dolphin, and beagle BLGs but probably not to manatee BLG [153,44].

The third group appears an odd one. All three contain the disulphide bridge Cys-66 -> Cys-160, and a free thiol - at positions 121, 50, 140 in BLG, OBP and MUP respectively. Their origins, from milk, nasal mucosa and mouse urine appear unrelated and the ligands they can bind show no obvious trends. OBP can bind a variety of ligands including retinol, whilst BLG can bind retinol

but not 2-isobutyl-3-methoxypyrazine, and MUP is believed to bind pheromones. If these proteins are closely related then the correlating facts remain undiscovered.

In addition to the five groups of this BLG superfamily, there is also another group of proteins which may be related. This new family contains smaller proteins typically about 12-14kDa, some of which are known to have 10-stranded β-barrel structures and can transport fatty acids or retinol - its members include FABPs and cellular retinoid binding proteins [246]. A summary of the properties of these two families is given in Figure 5:8, and shows that each has some distinct characteristics. It is tempting to suggest that the features of the new family are due to its members being cellular proteins, compared to those in the BLG superfamily which are not.

5:7 POSSIBLE FUNCTION OF BLG.

Although a lot more information is now available on the structure of, and binding to, BLG there still appears to be insufficient data to unequivocally assign a function to BLG. However one will be proposed, and the evidence for it examined.

5:7:1 PROPOSAL OF FUNCTION.

One possible role for BLG is as follows. BLG is secreted into the milk of mammals, where it could bind a free fatty acid in its external hydrophobic channel. This could induce a conformational change that enables the secondary binding site, the hydrophobic cavity, to become accessible for the binding of more FFAs. The BLG+FFA complex is then transferred from mother to young via suckling, and survives intact through the stomach. As

Figure 5:8

COMPARISON BETWEEN THE 8- AND 10-STRANDED β-BARREL FAMILIES.

Property	**8-Stranded family**	**10-Stranded family**
Members	BLG, RBP, APOD etc.	FABPs, cellular retinoid binding proteins, P2
Environment	secreted fluids	cellular
Molecular weight	18-20kDa	12-15kDa
Number of amino acids	160-180	about 130
Location of residues:		
G-W-	about residue 20	about residue 5
TDY--Y	about residue 100	not present
Disulphide bridges	two: one is 34-160 (APOD) or 66-160 (BLG,RBP)	none present
α-helix	residues 130-140 (BLG)	none near C-end
β-barrel	about $360A^3$ (BLG)	about $270A^3$ (flattened in IFABP)
Binding site for FFA	external channel (BLG)	β-barrel (IFABP)
Association constant	about 10^5mol/l (BLG)	about 10^8mol/l (FABP)
Number of coding exons	4 (APOD) to 7 (RBP)	4 (IFABP)

this complex progresses down the GI tract the pH increases and the Tanford transition may occur, altering the conformation of the region of the protein that is conserved in the members of the superfamily, and allowing the complex to bind to its specific receptors. A further conformational change may occur upon binding and enable the FFAs being carried in the β-barrel to be released.

5:7:2 EVIDENCE FOR THE PROPOSED FUNCTION.

The distribution of BLG amongst the milks of mammals is widespread. Its correlation with the placental type and method of immunoglobulin transport from mother to young may be fortuitous, as it has recently been suggested that a primate, macaque also has BLG in its milk [91]. This seems to suggest BLG is more wide-spread amongst mammals, and probably a requirement for the development of healthy young.

The large amount of BLG present in milk - it is the major whey protein present at about 3g/l in cows milk - suggests that it has a general function; whilst its large hydrophobic cavity($360A^3$), ability to bind a wide variety of ligands *in vitro*, and membership of a superfamily of transport proteins, implies that it could carry various hydrophobic ligands. The ability of BLG to bind FFAs in the same ratio present in milk [177], the recent isolation of BLGs with FFAs bound [240], and the higher concentration of fatty acids in those milks that have more BLG, eg. dolphin milk, indicates that BLG may be a general fatty acid transport protein.

The binding of palmitic acid to bovine BLG has been studied *in vitro* [176]. The ligand is believed to bind in the external hydrophobic channel with an association constant of $3.4x10^5$mol/l (compared to about 10^8mol/l for the binding of FFAs to cellular FABPs [242]), and induce a conformational change that enables the secondary

binding site, probably the hydrophobic cavity, to become available to up to another 6 palmitic acid molecules. There appears to be no reason why the same cannot occur in milk at pH around 6.6. It is unclear whether the BLG isolated by Aschaffenburg and Drewry [57] still contains any randomly orientated molecules in its β-barrel, but this seems likely, as a hexane extraction can release some FFAs [240].

Upon ingestion of milk by the young the complex passes through the highly acidic stomach. The protein survives intact as BLG is known to be stable down to pH 2, and in the presence of some proteases. However it is not clear whether the FFA in the external channel remains bound, although this seems unlikely. The dissociation that must have occurred upon reducing the pH is likely to affect the α-helical region of the protein which borders this binding site. However those molecules within the β-barrel are well protected from the acidic conditions, and likely to remain bound.

Upon progression along the GI tract the pH increases, and by the lower end is likely to have increased from 6 to 7.5 and induced the Tanford transition. This conformational change is localized to a small region of the protein which could be the area of BLG which binds to the BLG-specific receptors that have been detected in the lower third of 3-week old calf ileum [8]. Binding to the receptor may cause another conformational change and enable the release of the FFAs, and weaken the association of the protein to the receptor. The apo-protein may then be displaced by another 'loaded' one, prior to its absorption (6-9% is absorbed intact), and excretion in the urine - BLG has been observed in calf urine.

The role of BLG as a general fatty acid binding protein has yet to be confirmed, but seems plausible - considering the abundance of free fatty acids in milk,

its ability to bind them, its β-barrel motif which is similar to that of cellular FABPs, and the knowledge that it has specific receptors in calf ileum.

BIBLIOGRAPHY.

[1] A.H.Palmer : The preparation of a crystalline globulin from the albumin fraction of cow's milk.
J. Biol. Chem. **104** 359-372 (1934)

[2] H.A.McKenzie : β-lactoglobulins.
Milk Proteins : Chemistry and Molecular Biology **Vol II** 257-330 (Academic New York 1971)

[3] R.L.J.Lyster : Reviews of the progress of dairy science. Section C. Chemistry of milk proteins.
J. Dairy Res. **39** 279-318 (1972)

[4] R.Aschaffenburg, D.W.Green & R.M.Simmons : Crystal forms of β-lactoglobulin.
J. Mol. Biol. **13** 194-201 (1965)

[5] D.W.Green, R.Aschaffenburg, A.Camerman, J.C.Coppola, P.Dunnill, R.M.Simmons, E.S.Komorowski, L.Sawyer, E.M.C.Turner & K.F.Woods : Structure of bovine β-lactoglobulin at 6A resolution.
J. Mol. Biol. **131** 375-397 (1979)

[6] C.Tanford, L.G.Bunville & Y.Nozaki : The reversible transformation of β-lactoglobulin at pH 7.5.
J. Am. Chem. Soc. **81** 4032-4036 (1959)

[7] A.Sivaprasadarao & J.B.C.Findlay : The interaction of retinol-binding protein with its plasma-membrane receptor.
Biochem. J. **255** 561-569 (1988)

[8] M.Z.Papiz, L.Sawyer, E.E.Eliopoulos, A.C.T.North, J.B.C.Findlay, R.Sivaprasadarao, T.A.Jones, M.E. Newcomer & P.J.Kraulis : The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein.
Nature **324** 383-385 (1986)

[9] L.Sawyer : One fold among many.
Nature **327** 659 only (1987)

[10] B.L.Larson : Biosynthesis and secretion of milk proteins: a review.

J.Dairy Sci. **46** 161-174 (1979)

[11] M.Yoshikawa, T.Mizukami, R.Sasaki & H.Chiba : Pre-β-lactoglobulin synthesis by mRNA from bovine mammary gland.
Agri. Biol. Chem. **42** 2185-2186 (1978)

[12] J.-C. Mercier & P.Gaye : Early events in secretion of main milk proteins: occurrence of precursors.
J.Dairy Sci. **65** 299-316 (1982)

[13] K.Bell, H.A.McKenzie, W.H.Murphy & D.C.Shaw : β-lactoglobulin droughtmaster: a unique protein variant.
Biochem. Biophys. Acta **214** 427-436 (1970)

[14] H.A.McKenzie, V.J.Muller & G.B.Treacy : "Whey" proteins of milk of the red (Macropus rufus) and Eastern grey (Macropus giganteus) kangaroo.
Comp. Biochem. Physiol. **74B** 259-271 (1983)

[15] J.Godovac-Zimmermann & D.C.Shaw : The primary structure, binding site and possible function of β-lactoglobulin from Eastern grey kangaroo (Macropus giganteus).
Biol. Chem. Hoppe-Seyler **368** 879-886 (1987)

[16] E.Kessler & K.Brew : Personal Communication cited in reference [41].

[17] R.L.J.Lyster, R.Jenness, N.I.Phillips & R.E.Sloan : Comparative biochemical studies of milks - III. Immuno-electrophoretic comparisons of milk proteins of the Artiodactyla.
Comp. Biochem. Physiol. **17** 967-971 (1966)

[18] J.Liberatori, L.Morisio Guidetti, A.Conti & L. Napolitano : β-lactoglobulins in the mammary secretions of camel (Camelus dromedarius) and she-ass. Immunological detection and preliminary physico-chemical characterization.
Boll. Soc. It. Biol. Sper. **LV** 1369-1373 (1979)

[19] K.Bell & H.A.McKenzie : β-lactoglobulins.
Nature **204** 1275-1279 (1964)

[20a] J.Liberatori, L.Morisio Guidetti & A.Conti : Immunological evidence of β-lactoglobulins in human colostrum and milk.
Boll. Soc. It. Biol. Sper. **LV** 822-825 (1979)

[20b] A.Conti, J.Liberatori & L.Napolitano : Isolation and preliminary physico-chemical characterization of human β-lactoglobulin.
Milchwissenschaft **35** 65-68 (1980)

[21] A.Conti, L.Napolitano & J.Liberatori : Characterization of β-lactoglobulins from human colostrum. Some physico-chemical properties and preliminary amino acid evaluation.
Dairy Industries International **47** 163-164 (1982)

[22a] G.Brignon, A.Chtourou & B.Ribadeau-Dumas : Does β-lactoglobulin occur in human milk?
J. Dairy Res. **52** 249-254 (1985)

[22b] J.C.Monti,A.-F.Mermoud & P.Jolles : Anti-bovine β-lactoglobulin antibodies react with a human lactoferrin fragment and bovine β-lactoglobulin present in human milk.
Experientia **45** 178-181 (1989)

[23] I.Axelsson, I.Jakobsson, T.Lindberg & B.Benediktsson : Bovine β-lactoglobulin in the human milk.
Acta Paed. Scand. **75** 702-707 (1986)

[24] S.G.Hambling, A.McAlpine & L.Sawyer : β-lactoglobulin : a review.
Chapter in Advanced Dairy Chemistry **Vol I** (in press)

[25] S.Ali & A.J.Clark : Personal Communication (1988)

[26] L.G.Hennighausen & A.E.Sippel : Mouse whey acidic protein is a novel member of the family of 'four-disulphide core' proteins.
Nucl. Acid Res. **10** 2677-2684 (1982)

[27] D.H.Lockwood, R.W.Turkington & Y.J.Topper : Hormone-dependent development of milk protein synthesis in mammary gland <u>in vitro</u>.
Biochim. Biophys. Acta **130** 493-501 (1966)

[28] L.Sawyer & C.Holt : Personal Communication (1888)

[29] A.J.Clark, P.Simons, I.Wilmut & R.Lathe : Pharmaceuticals from transgenic livestock. TIBTech. **5** 20-24 (1987)

[30] J.P.Simons, M.McClenaghan & A.J.Clark : Alteration of the quality of milk by expression of sheep β-lactoglobulin in transgenic mice. Nature **328** 530-532 (1987)

[31] D.Marcon-Genty, D.Tome, O.Kheroua, A.-M.Dumontier, M.Heyman & J.-F.Desjeux : Transport of β-lactoglobulin across rabbit ileum in vitro. Am. J. Physiol. **256** G943-G948 (1989)

[32] F.Grosclaude, M.-F.Mahe, J.-C.Mercier, J.Bonnemarie & J.H.Teissier : Polymorphisme des lactoproteines de bovines nepalais. Ann. Genet. Sel. anim **8** 461-479 (1976)

[33] A.Conti & J.Godovac-Zimmermann : Antibodies raised against bovine β-lactoglobulin react with β_2-microglobulin. Biol. Chem. Hoppe-Seyler **371** 261-263 (1990)

[34] T.Nagasawa, I.Kiyosawa, R.Kato & K.Kuwahara : Isolation of canine α_s-casein and major whey protein component A and their composition. J. Dairy Sci. **55** 1550-1556 (1972)

[35] B.A.Askonas : Crystallization of goat β-lactoglobulin. Biochem. J. **58** 332-336 (1954)

[36] J.L.Maubois, R.Pion & B.Ribadeau-Dumas : Preparation et etude de la β-lactoglobuline de brebis crystallisee. Biochim. Biophys. Acta **107** 501-510 (1965)

[37] J.Godovac-Zimmermann, A.Conti & L.Napolitano : The complete amino-acid sequence of dimeric β-lactoglobulin from mouflon (Ovis ammon musimon) milk. Biol. Chem. Hoppe-Seyler **368** 1313-1319 (1987)

[38] E.I.McDougall & J.C.Stewart : The whey proteins of the milk of red deer (Cervus elaphus L.).

Biochem. J. **153** 647-655 (1976)

[39] C.Holt : Personal Communication (1988).

[40] P.N.Shubin, M.N.Turubanov & V.S.Matyukov : An electrophoretic study of milk proteins in reindeer and European elk.
Zh. Obshch. Biol. **32** 746-750 (1971)

[41] E.Kessler & K.Brew : The whey proteins of pig's milk : isolation and characterization of a β-lactoglobulin.
Biochim. Biophys. Acta **200** 449-458 (1970)

[42] K.Bell, H.A.McKenzie, V.Muller, C.Rogers & D.C.Shaw: Equine whey proteins.
Comp. Biochem. Physiol. **68B** 225-236 (1981)

[43] J.Godovac-Zimmermann, A.Conti, L.James & L.Napolitano : Micro-analysis of the amino-acid sequence of monomeric β-lactoglobulin I from donkey (Equus asinus) milk.
Biol. Chem. Hoppe-Seyler **369** 171-179 (1988)

[44] S.Pervaiz & K.Brew : Purification and characterization of the major whey proteins from the milks of the Bottlenose dolphin (Tursiops truncatus), the Florida manatee (Trichechus manatus latirostris) and the beagle (Canis familiaris).
Arch. Biochem. Biophys. **246** 846-854 (1986)

[45] O.M.R.Westwood, M.G.Chapman, N.Totty, R.Philp, A.E.Bolton & N.R.Lazarus : N-terminal sequence analysis of human placental protein 14 purified in high yield from decidual cytosol.
J. Reprod. Fert. **82** 493-500 (1988)

[46] U.S.Ashworth, G.D.Ramaiah & M.C.Keyes : Species difference in the composition of milk with special reference to the Northern fur seal.
J. Dairy. Sci. **49** 1206-1211 (1966)

[47] K.Brew & P.N.Campbell : The characterization of the whey proteins of guinea-pig milk.
Biochem. J. **102** 258-264 (1967)

[48] F.M.Fernandez & G.Oliver : Proteins present in llama milk. I. Quantitative aspects and general characteristics.
Milchwissenschaft **43** 299-302 (1988)

[49] J.A.Halliday, K.Bell, H.A.McKenzie & D.C.Shaw : Feline whey proteins : identification, isolation and initial characterization of α-lactalbumin, β-lactoglobulin and lysozyme.
Comp. Biochem. Physiol. **95B** 773-779 (1990)

[50] H.Mossmann : Vertebrate fetal membranes (1987)

[51] J.E.Butler : Immunoglobulins of mammary secretions.
Lactation: A comprehensive treatise
Vol **III** 217-256 (1974)

[52] M.Julkunen, M.Seppala & O.A.Janne : Complete amino acid sequence of human placental protein 14: a progesterone-regulated uterine protein homologous to β-lactoglobulins.
Proc. Natl. Acad. Sci. **85** 8845-8849 (1988)

[53] M.Seppala, M.-L.Huhtala, M.Julkunen, R.Koistinen & E.-M.Rutanen : Uterine proteins, nomenclature determined by biological action.
Res. in Reprod. **19** 2 only (1987)

[54] S.C.Bell : Purification of human secretory pregnancy-associated endometrial α_2-globulin (α_2-PEG) from cytosol of first trimester pregnancy endometrium.
Human Reprod. **1** 313-318 (1986)

[55] M.-L.Huhtala, M.Seppala, A.Narvanen, P.Palomaki, M.Julkunen & H.Bohn : Amino acid sequence homology between human placental protein 14 and β-lacto-globulins from various species.
Endocrinol. **120** 2620-2622 (1987)

[56] S.C.Bell, J.W.Keyte & G.T.Waites : Pregnancy-associated endometrial α_2-globulin (α_2-PEG), the major secretory protein of the luteal phase and first trimester pregnancy endometrium, is not a

glycosylated prolactin but related to β-lactoglobulins.
J. Clin. Endocrinol. Metab. **65** 1067-1071 (1987)

[57] R.Aschaffenburg & J.Drewry : Improved method for the preparation of crystalline β-lactoglobulin and α-lactalbumin from cow's milk.
Biochem. J. **65** 273-277 (1957)

[58] J.McD.Armstrong, H.A.McKenzie & W.H.Sawyer : On the fractionation of β-lactoglobulin and α-lactalbumin.
Biochim. Biophys. Acta **147** 60-72 (1967)

[59] H.L.Monaco, G.Zanotti, P.Spadon, M.Bolognesi, L.Sawyer & E.E.Eliopoulos : Crystal structure of the trigonal form of bovine β-lactoglobulin and its complex with retinol at 2.5A resolution.
J. Mol. Biol. **197** 695-706 (1987)

[60] G.Braunitzer, J.Liberatori & H.-J.Kolde : The primary structure of the β-lactoglobulin of the waterbuffalo (Bubalus arnee).
Z. Naturforsch. **34c** 880-881 (1979)

[61] E.B.Kalan & J.J.Basch : Preparation of goat β-lactoglobulin.
J. Dairy Sci. **49** 406-409 (1969)

[62] G.Preaux, G.Braunitzer, B.Schrank & A.Stangl : The amino acid sequence of goat β-lactoglobulin.
Hoppe-Seyler Z. Physiol. Chem. **360** 1595-1604 (1979)

[63] G.Preaux, G.Braunitzer & H.-J.Kolde : Primary structure of ovine β-lactoglobulin.
Arch. Int. Physiol. Biochem. **88** B45 35-36 (1980)

[64] J.Godovac-Zimmermann, A.Conti, J.Liberatori & G. Braunitzer : The amino-acid sequence of β-lactoglobulin II from horse colostrum (Equus caballus, Perissodactyla) : β-lactoglobulins are retinol-binding proteins.
Biol. Chem. Hoppe-Seyler **366** 601-608 (1985)

[65] E.B.Kalan, R.R.Kraeling & R.J.Gerrits : Isolation and partial characterization of a polymorphic swine whey protein.

Int. J. Biochem. **2** 232-244 (1971)

[66] K.Bell, H.A.McKenzie & D.C.Shaw : Porcine β-lactoglobulin A and C.
Mol. Cell. Biochem. **35** 103-111 (1981)

[67] I.Krause, J.Buchberger, G.Weiss, M.Pflugler & H.Klostermeyer : Isoelectric focusing in immobilized pH gradients with carrier ampholytes added for high-resolution phenotyping of bovine β-lactoglobulins: characterization of a new genetic variant.
Electrophoresis **9** 609-613 (1988)

[68] G.Erhardt, J.Godovac-Zimmermann & A.Conti : Isolation and complete primary sequence of a new ovine wild-type β-lactoglobulin C.
Biol. Chem. Hoppe-Seyler **370** 757-762 (1989)

[69] G.Brignon & B.Ribadeau-Dumas : Localisation dans la chaine peptidique de la β-lactoglobuline bovine de la substitution Glu/Gln differenciant les variants genetiques B et D.
FEBS Lett. **33** 73-76 (1973)

[70] K.A.Piez, E.W.Davie, J.E.Folk & J.A.Gladner : β-lactoglobulins A and B.
J. Biol. Chem. **236** 2912-2916 (1961)

[71] G.Braunitzer, R.Chen, B.Schrank & A.Stangl : Die sequenzanalyse des β-lactoglobulins.
Hoppe-Seyler Z. Physiol. Chem. **354** 867-878 (1973)

[72a] K.Bell & H.A.McKenzie : The isolation and properties of bovine β-lactoglobulin C.
Biochim. Biophys. Acta **147** 109-122 (1967)

[72b] H.A.McKenzie & D.C.Shaw : Personal Communication cited in [75]

[73] D.R.Osterhoff & A.M.G.Pretorius : Inherited biochemical polymorphism in milk proteins.
Proc. S. Afr. Soc. Anim. Prod. **5** 166-173 (1966)

[74] G.Brignon, B.Ribadeau-Dumas, J.Garnier, D.Pantaloni, S.Guinard, J.J.Basch & S.N.Timasheff : Chemical and

physico-chemical characterization of genetic variant D of bovine β-lactoglobulin.
Arch. Biochem. Biophys. **129** 720-727 (1969)

[75] K.Bell, H.A.McKenzie & D.C.Shaw : Bovine β-lactoglobulin E, F and G of Bali (Banteng) Cattle, Bos (Bibos) javanicus.
Aust. J. Biol. Sci. **34** 133-147 (1981)

[76] A.Conti, L.Napolitano, A.M.Cantisani, R.Davoli & S.Dall'Olio : Bovine β-lactoglobulin H : isolation by preparative isoelectric focusing in immobilized pH gradients and preliminary characterization.
J. Biochem. Biophys. Meth. **16** 205-214 (1988)

[77] K.Bell & H.A.McKenzie : The whey proteins of ovine milk : β-lactoglobulins A and B.
Biochim. Biophys. Acta **147** 123-134 (1967)

[78] H.-J.Kolde & G.Braunitzer : The primary structure of ovine β-lactoglobulin. 2. Discussion and genetic aspects.
Milchwissenschaft **38** 70-72 (1983)

[79] A.Sen & S.Chaudhuri : Non-casein proteins of goat's milk.
Nature **195** 286-287 (1962)

[80] S.Pervaiz & K.Brew : Homology of β-lactoglobulin, serum retinol-binding protein, and protein HC.
Science **228** 335-337 (1985)

[81a] A.Conti, J.Liberatori & L.Napolitano : Heterogeneity of equine β-lactoglobulins.
Milchwissenschaft **37** 338-340 (1982)

[81b] A.Conti, J.Godovac-Zimmermann, J.Liberatori & G.Braunitzer : The primary structure of monomeric β-lactoglobulin I from horse colostrum (Equus caballus, Perissodactyla).
Hoppe-Seyler Z. Physiol. Chem. **366** 1393-1401 (1984)

[82] R.R.Kraeling & R.J.Gerrits : Polymorphism of a protein of sows' whey.
J. Dairy Sci. **52** 2036-2038 (1969)

[83a] A.Conti, L.Napolitano, J.Godovac Zimmermann, J.Liberatori & G.Braunitzer : Preparative separation of pig polymorphic β-lactoglobulins by IPG and non-IPG.
Protides Biol. Fluids Proc. Colloq. **34** 907-910 (1986)

[83b] J.Liberatori, L.Morisio Guidetti & A.Conti : Immunochemical studies on β-lactoglobulins. Precipitin reactions of sow's and mare's mammary secretions against anti-bovine β-lactoglobulin antiserum.
Boll. Soc. It. Biol. Sper. **LV** 815-821 (1979)

[84] A.Conti, J.Godovac-Zimmermann, F.Pirchner, J.Liberatori & G.Braunitzer : Pig β-lactoglobulin I (Sus scrofa domestica, Artiodactyla).
Biol. Chem. Hoppe-Seyler **367** 871-878 (1986)

[85] J.Godovac-Zimmermann & G.Braunitzer : Modern aspects of the primary structure and function of β-lactoglobulins.
Milchwissenschaft **42** 294-297 (1987)

[86] J.Godovac-Zimmermann, I.Krause, J.Buchberger, G.Weiss & H.Klostermeyer : A novel wild-type of β-lactoglobulin W, and its primary sequence.
Biol. Chem. Hoppe-Seyler **371** 255-260 (1990)

[87] H.B.Bull & B.T.Currie : Osmotic pressure of β-lactoglobulin solutions.
J. Am. Chem. Soc. **68** 742-745 (1946)

[88] R.Cecil & A.G.Ogston : The sedimentation constant, diffusion constant and molecular weight of lactoglobulin.
Biochem. **44** 33-35 (1949)

[89] M.Halwer, G.C.Nutting & B.A.Brice : Molecular weight of lactoglobulin, ovalbumin, lysozyme and serum albumin by light scattering.
J. Am. Chem. Soc. **73** 2786-2790 (1951)

[90] H.B.Bull : Monolayers of β-lactoglobulin. II. Film molecular weight.

J. Am. Chem. Soc. **68** 745-747 (1946)

[91] D.C.Shaw : Personal Communication (1990)

[92] R.Townend, L.Weinberger & S.N.Timasheff : Molecular interactions in β-lactoglobulin. IV. The dissociation of β-lactoglobulin below pH 3.5.
J. Am. Chem. Soc. **82** 3175-3179 (1960)

[93] G.Patterson & L.A.Fothergill-Gilmore : Personal Communication (1990)

[94] S.N.Timasheff & R.Townend : Molecular interactions in B-lactoglobulin. V. The association of the genetic species of β-lactoglobulin below the isoelectric point.
J. Am. Chem. Soc. **83** 464-469 (1961)

[95] C.Georges & S.Guinard : Sur la dissociation reversible de la β_2-lactoglobuline a des pH superieurs a 5.5. I. Etude par la diffusion de la lumiere.
J. Chim. Phys. **57** 606-614 (1960)

[96] Unused.

[97] H.Pessen, J.M.Purcell & H.M.Farrell(Jr) : Proton relaxation rates of water in dilute solutions of β-lactoglobulin. Determination of cross relaxation and correlation with structural changes by the use of two genetic variants of a self-associating globular protein.
Biochim. Biophys. Acta **828** 1-12 (1985)

[98] R.Townend & S.N.Timasheff : Molecular interactions in β-lactoglobulin III. Light scattering investigation of the stoichiometry of the association between pH 3.7 and 5.2.
J. Am. Chem. Soc. **82** 3168-3174 (1960)

[99] S.N.Timasheff & R.Townend : Structure of the β-lactoglobulin tetramer.
Nature **203** 517-519 (1964)

[100] J.McD.Armstrong & H.A.McKenzie : A method for modification of carboxyl groups in proteins: its application to the association of bovine

β-lactoglobulin A.
Biochim. Biophys. Acta **147** 93-99 (1967)

[101] E.I.McDougall & J.C.Stewart : The state of aggregation of red deer (Cervus elaphus L.) β-lactoglobulin preparations near neutral pH.
Biochem. J. **167** 45-51 (1977)

[102] C.Georges, S.Guinand & J.Tonnelat : Etude thermodynamique de la dissociation reversible de la β-lactoglobuline B pour des pH superieurs a 5.5.
Biochim. Biophys. Acta **59** 737-739 (1962)

[103] H.A.McKenzie & W.H.Sawyer : The effect of pH on β-lactoglobulins.
Nature **214** 1101-1104 (1967)

[104] S.N.Timasheff, L.Mescanti, J.J.Basch & R.Townend : Conformational transitions of bovine β-lacto-globulins A, B, and C.
J. Biol. Chem. **241** 2496-2501 (1966)

[105] M.L.Groves, N.J.Hipp & T.L.McMeekin : Effect of pH on the denaturation of β-lactoglobulin and its dodecyl sulphate derivative.
J. Am. Chem. Soc. **73** 2790-2793 (1951)

[106] H.Roels, G.Preaux & R.Lontie : Stabilization of β-lactogloblin A and B at pH 8.9 by blocking the thiol groups.
Arch. Int. Biochem. **74** 522-523 (1966)

[107] Unused.

[108] Unused.

[109] S.N.Timasheff, R.Townend & L.Mescanti : The optical rotatory dispersion of β-lactoglobulins.
J. Biol. Chem. **241** 1863-1870 (1966)

[110] J.M.Purcell & H.Susi : Solvent denaturation of proteins as observed by resolution-enhanced Fourier transform infrared spectroscopy.
J. Biochem. Biophys. Meth. **9** 193-199 (1984)

[111] B.G.Frushour & J.L Koenig : Raman studies of the crystalline, solution and alkaline-denatured

states of β-lactoglobulin.
Biopolymers **14** 649-662 (1975)

[112] K.H.Park & D.B.Lund : Calorimetric study of thermal denaturation of β-lactoglobulin.
J. Dairy. Sci **67** 1699-1706 (1984)

[113] B.L.Larson & R.Jenness : Characterization of the sulphydryl groups and the kinetics of the heat denaturation of crystalline β-lactoglobulin.
J. Am. Chem. Soc. **74** 3090-3093 (1952)

[114] O.E.Mills : Effect of temperature on tryptophan fluorescence of β-lactoglobulin B.
Biochim. Biophys. Acta **434** 324-332 (1976)

[115] M.Dupont : Etude d'une etape reversible dans la thermo-denaturation de la β-lactoglobuline bovine A.
Biochim. Biophys. Acta **102** 500-513 (1965)

[116] H.Otani, S.-I.Morita & F.Tokita : Studies on the antigenicity of the browning product between β-lactoglobulin and lactose.
Jap. J. Zootech. Sci. **56** 67-74 (1985)

[117] W.H.Sawyer : Complex between β-lactoglobulin and κ-casein. A review.
J. Dairy Sci **52** 1347-1355 (1969)

[118] R.F.Greene & C.N.Pace : Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin and β-lactoglobulin.
J. Biol. Chem. **249** 5388-5393 (1974)

[119] H.A.McKenzie & G.B.Ralston : The denaturation of proteins : two states? reversible or irreversible?
Experientia **27** 617-624 (1971)

[120] G.Preaux & H.Dekoster : Fluorescence of β-lactoglobulin B as a function of pH before and after blocking the thiol groups with organomercurials.
Hoppe-Seyler Z. Physiol. Chem. **355** 1238-1239 (1974)

[121] Unused.

[122] Unused.

[123] C.Tanford & P.K.De : The unfolding of β-lactoglobulin at pH 3 by urea, formamide, and other organic substances.
J. Biol. Chem. **236** 1711-1715 (1961)

[124] A.J.Hillquist Damon & G.C.Kresheck : Influence of surfactants on the conformation of β-lactoglobulin B using circular dichroism.
Biopolymers **21** 895-908 (1982)

[125] P.Dunnill, D.W.Green & R.M.Simmons : Heavy-atom sulphydryl derivatives of ox haemoglobin and β-lactoglobulin; factors affecting isomorphism of native and derivative crystals.
J. Mol. Biol. **22** 135-144 (1966)

[126] D.Pantaloni : Structure et changements de conformations de la β-lactoglobuline en solution.
Thesis for doctorate, University of Paris (1965)

[127] R.Lontie & G.Preaux : Polarimetric investigation of β-lactoglobulin A and B and of the reactivity of their thiol groups.
Protides Biol. Fluids **14** 475-482 (1966)

[128] D.Pantaloni : Etude de la transition R-S de la β-lactoglobuline par spectropolarimetrie et par spectrophotometrie de differences.
Compte. Rend. Seanc. Acad. Sci. Paris **258** 5753-5756 (1964)

[129] D.Pantaloni : Role des groupes SH dans la structure de la β-lactoglobuline.
Compte. Rend. Acad. Sci. **254** 1884-1886 (1962)

[130] Unused.

[131] Unused.

[132] R.Townend, T.F.Kumosinski & S.N.Timasheff : The circular dichroism of variants of β-lactoglobulin.
J. Biol. Chem. **242** 4538-4545 (1967)

[133] S.N.Timasheff & H.Susi : Infrared investigation of the secondary structure of β-lactoglobulins.
J. Biol. Chem. **241** 249-251 (1966)

[134] H.Deckmyn & G.Preaux : Chain folding prediction of the bovine β-lactoglobulins.
Arch. Int. Physiol. Biochim. **86** 938-939 (1978)

[135] L.K.Creamer, D.A.D.Parry & G.N.Malcolm : Secondary structure of bovine β-lactoglobulin B.
Arch. Biochem. Biophys. **227** 98-105 (1983)

[136] Unused.

[137] C.Tanford & V.G.Taggart : Ionization-linked changes in protein conformation. II. The N-R transition in β-lactoglobulin.
J. Am. Chem. Soc. **83** 1634-1638 (1961)

[138] P.Dunnill & D.W.Green : Sulphydryl groups and the N-R conformational change in β-lactoglobulin.
J. Mol. Biol. **15** 147-151 (1965)

[139] R.L.J.Lyster : Reactivity of the -SH groups of β-lactoglobulin subunits.
Abst. 6th Int. Congress Biochem. (1964)

[140] L.Sawyer & D.W.Green : The reaction of cow β-lactoglobulin with tetracyanoaurate(III).
Biochim. Biophys. Acta **579** 234-239 (1979)

[141] S.J.Yewdall : Structural studies on β-lactoglobulin.
Ph.D. thesis, University of Leeds (1988)

[142] F.Mainferme, G.Preaux & R.Lontie : Location of the disulphide bridges in the sequence of β-lactoglobulin A with CNBr and thermolysin.
Arch. Int. Physiol. Biochem. **79** 840-841 (1971)

[143] J.Martial, G.Preaux & R.Lontie : Location of the thiol group and of the disulphide bridges in the sequence of β-lactoglobulin B.
Arch. Int. Physiol. Biochem. **79** 842-843 (1971)

[144] G.Perez Gomez, G.Preaux & R.Lontie : Location of the disulphide bridges in β-lactoglobulin B with pepsin and trypsin after CNBr cleavage.
Arch. Int. Physiol. Biochem. **79** 843-844 (1971)

[145] D.M.Byler, H.Susi & H.M.Farrell(Jr) : Laser-Raman spectra, sulphydryl groups, and conformation of

the cystine linkages of β-lactoglobulin.
Biopolymers **22** 2507-2511 (1983)

[146] H.A.McKenzie & D.C.Shaw : Alternative positions for the sulphydryl group in β-lactoglobulin : the significance for sulphydryl location in other proteins.
Nature **238** 147-149 (1972)

[147] H.A.McKenzie, G.B.Ralston & D.C.Shaw : Location of sulphydryl and disulphide groups in bovine β-lactoglobulins and effects of urea.
Biochemistry **11** 4539-4547 (1972)

[148] Unused.

[149] W.L.Stone & A.Wishnia : Binding of iodomercurates to sulphydryl-blocked β-lactoglobulin -A, -B, and -C.
Bioinorg. Chem. **8** 517-529 (1978)

[150] R.Townend, T.T.Herskovits, S.N.Timasheff & M.J.Gorbunoff : The state of amino acid residues in β-lactoglobulin.
Arch. Biochem. Biophys. **129** 567-580 (1969)

[151] Unused.

[152] W.A.Prutz, J.Butler, E.J.Land & A.J.Swallow : Direct demonstration of electron transfer between tryptophan and tyrosine in proteins.
Biochem. Biophys. Res. Comm. **96** 408-414 (1980)

[153] R.D.Fugate & P.-S.Song : Spectroscopic characterization of β-lactoglobulin-retinol complex.
Biochim. Biophys. Acta **625** 28-42 (1980)

[154] S.L.Woo, L.K.Creamer & T.Richardson : Chemical phosphorylation of bovine β-lactoglobulin.
J. Agri. Food Chem. **30** 65-70 (1982)

[155] E.M.Brown, P.E.Pfeffer, T.F.Kumosinski & R.Greenberg : Accessibility and mobility of lysine residues in β-lactoglobulin.
Biochem. **27** 5601-5610 (1988)

[156] R.Greenberg & E.B.Kalan : Studies on β-lactoglobulins A,B and C. II. Preparation of modified proteins by treatment with carboxypeptidase A.
Biochemistry **4** 1660-1666 (1965)

[157] M.I.Halpin & T.Richardson : Selected functionality changes of β-lactoglobulin upon esterification of side-chain carboxyl groups.
J. Dairy. Sci. **68** 3189-3198 (1985)

[158] L.Sawyer, M.Z.Papiz, A.C.T.North & E.E.Eliopoulos : Structure and function of bovine β-lactoglobulin.
Biochem. Soc. Trans. **13** 265-266 (1985)

[159] J.D.Garrick : Is the role of β-lactoglobulin to serve as a retinol (vitamin A) carrier in milk?
Thesis for BSc(Hons), University of Aberdeen (1986)

[160] R.Hemley, B.E.Kohler & P.Siviski : Absorption spectra for the complexes formed from vitamin-A and β-lactoglobulin.
Biophys. J. **28** 447-455 (1979)

[161] S.Futtermann & J.Heller : The enhancement of fluorescence and the decreased susceptibility to enzymatic oxidation of retinol complexed with bovine serum albumin, β-lactoglobulin and the retinol-binding protein of human plasma.
J. Biol. Chem. **247** 5168-5172 (1972)

[162] R.Lovrien & W.F.Anderson : Resolution of binding sites in β-lactoglobulin.
Arch. Biochem. Biophys. **131** 139-144 (1969)

[163] M.P.Thompson & H.M.Farrell(Jr) : Genetic variants of the milk proteins.
Lactation. A Comprehensive Treatise
Vol **III** 109-134 (1974)

[164] H.M.Farrell(Jr), M.J.Behe & J.A.Enyeart : Binding of p-nitrophenyl phosphate and other aromatic compounds by β-lactoglobulin.
J. Dairy Sci. **70** 252-258 (1987)

[165] Unused.

[166] Unused.

[167] K.A.Robillard(Jr) & A.Wishnia : Aromatic hydrophobes and β-lactoglobulin A. Thermodynamics of binding.
Biochemistry **11** 3835-3840 (1972)

[168] A.Mohammadzadeh-K., R.E.Feeney & L.M.Smith : Hydrophobic binding of hydrocarbons by proteins. I. Relationship of hydrocarbon structure.
Biochim. Biophys. Acta **194** 246-255 (1969)

[169] A.Wishnia & T.W.Pinder(Jr) : Hydrophobic interactions in proteins. The alkane binding site of β-lactoglobulins A and B.
Biochemistry **5** 1534-1542 (1966)

[170] T.E.O'Neill & J.E.Kinsella : Binding of alkanone flavors to β-lactoglobulin: effects of conformational and chemical modification.
J. Agri. Food Chem. **35** 770-774 (1987)

[171] Unused.

[172] Unused.

[173] T.S.Seibles : Interaction of dodecyl sulfate with native and modified β-lactoglobulin.
Biochemistry **8** 2949-2954 (1969)

[174] T.L.McMeekin, B.D.Polis, E.S.Della Monica & J.H.Custer : A crystalline complex of β-lactoglobulin with dodecyl sulfate.
J. Am. Chem. Soc. **71** 3606-3609 (1949)

[175] Y.-Y.T.Su & B.Jirgensons : Further studies on detergent-induced conformational transitions in proteins. Circular dichroism of ovalbumin, bacterial α-amylase, papain, and β-lactoglobulin at various pH values.
Arch. Biochem. Biophys. **181** 137-146 (1977)

[176] A.A.Spector & J.E.Fletcher : Binding of long chain fatty acids to β-lactoglobulin.
Lipids **5** 403-411 (1970)

[177] M.C.Diaz de Villegas, R.Oria, F.J.Sala & M.Calvo :

Lipid binding by β-lactoglobulin of cow milk.
Milchwissenschaft **42** 357-358 (1987)

[178] E.O.Keith : Personal Communication (1987)

[179] L.M.Smith, P.Fantozzi & R.K.Creveling : Study of triglyceride-protein interaction using a microemulsion-filtration method.
J. Am. Oil Chem. Soc. **60** 960-967 (1983)

[180] E.M.Brown, R.J.Carroll, P.E.Pfeffer & J.Sampugna : Complex formation in sonicated mixtures of β-lactoglobulin and phosphatidylcholine.
Lipids **18** 111-118 (1983)

[181] H.P.Baker & H.A.Saroff : Binding of sodium ions to β-lactoglobulin.
Biochemistry **4** 1670-1677 (1965)

[182] L.W.Cunningham & B.J.Nuenke : Analysis of modified β-lactoglobulins and ovalbumins prepared from the sulfenyl iodide intermediates.
J. Biol. Chem. **235** 1711-1715 (1960)

[183] E.M.Brown & H.M.Farrell(Jr) : Interaction of β-lactoglobulin and cytochrome c: complex formation and iron reduction.
Arch. Biochem. Biophys. **185** 156-164 (1978)

[184] R.J.Workman, M.M.McKown & R.I.Gregerman : Renin: inhibition by proteins and peptides.
Biochemistry **13** 3029-3035 (1974)

[185] H.G.Hunziker & N.P.Tarassuk : Chromatographic evidence for heat-induced interaction of α-lactalbumin and β-lactoglobulin.
J. Dairy Sci. **48** 733-734 (1965)

[186] J.J.Baumy & G.Brule : Binding of bivalent cations to α-lactalbumin and β-lactoglobulin: effect of pH and ionic strength.
Le Lait **68** 33-48 (1988)

[187] F.Lampreave, A.Pineiro, J.H.Brock, H.Castillo, L.Sanchez & M.Calvo : Interaction of bovine lactoferrin with other proteins of milk whey.

Int. J. Biol. Macromol. **12** 2-5 (1990)

[188] Z.Haque, M.M.Kristjansson & J.E.Kinsella : Interaction between κ-casein and β-lactoglobulin: possible mechanism.
J. Agri. Food Chem. **35** 644-649 (1987)

[189] R.Jenness, N.I.Phillips & E.B.Kalan : Immuno-chemical comparison of β-lactoglobulins.
Fed. Proc. **26** 340 only (1967)

[190] F.S.Mathews : The structure, function and evolution of cytochromes.
Prog. Biophys. Molec. Biol. **45** 1-56 (1985)

[191] Unused.

[192] Unused.

[193] E.Lebenthal, J.Laor, Z.Lewitus, Y.Matoth & S.Frier : Gastrointestinal protein loss in allergy to cows milk β-lactoglobulin.
Israel J. Med. Sci. **6** 506-510 (1970)

[194] S.P.Fallstrom, S.Ahlstedt & L.A.Hanson : Specific antibodies in infants with gastrointestinal intolerance to cow's milk protein.
Int. Archs. Allergy appl. Immun. **56** 97-105 (1978)

[195] Unused.

[196] Unused.

[197] D.M.Roberton, R.Paganelli, R.Dinwiddie & R.J.Levinsky : Milk antigen absorption in the preterm and term neonate.
Arch. Dis. Child. **57** 369-372 (1982)

[198] Unused.

[199] M.Frick & C.H.L.Rieger : Local antibodies to α-casein and β-lactoglobulin in the saliva of infants.
Pediat. Res. **22** 399-401 (1987)

[200] T.N.Koritz, S.Suzuki & R.R.A.Coombs : Antigenic stimulation with proteins of cow's milk via the oral route in guinea pigs and rats. 1. Measurement of antigenically intact β-lactoglobulin and casein

in the gastrointestinal contents of duodenum, jejunum and ileum.
Int. Archs. Allergy appl. Immun. **82** 72-75 (1987)

[201] S.Suzuki, T.N.Koritz & R.R.Coombs : Antigenic stimulation with proteins of cow's milk via the oral route in guinea pigs and rats. 2. Antibodies to β-lactoglobulin secreted into the alimentary canal and serum.
Int. Archs. Allergy appl. Immun. **82** 76-82 (1987)

[202] D.A.Granato & P.F.Piguet : A mouse monoclonal IgE antibody anti bovine milk β-lactoglobulin allows studies of allergy in the gastrointestinal tract.
Clin. exp. Immunol. **63** 703-710 (1986)

[203] Z.Malik, R.Bottomley & B.Austen : Allergenic properties of the genetic variants A and B of bovine β-lactoglobulin.
Int. Archs. Allergy appl. Immun. **86** 245-248 (1988)

[204] Unused.

[205] J.Kurisaki, S.Nakamura, S.Kaminogawa & K.Yamauchi : The antigenic properties of β-lactoglobulin examined with mouse IgE antibody.
Agri. Biol. Chem. **46** 2069-2075 (1982)

[206] Q.Huang, J.W.Coleman & D.R.Stanworth : Investigation of the allergenicity of β-lactoglobulin and its cleavage fragments.
Int. Archs. Allergy appl. Immun. **78** 337-344 (1985)

[207] Unused.

[208] J.Kurisaki, S.Nakamura, S.Kaminogawa, K.Yamauchi, S.Watanabe, K.Hotta & M.Hattori : Antigenicity of modified β-lactoglobulin examined by three different assays.
Agri. Biol. Chem. **49** 1733-1737 (1985)

[209] H.Otani, T.Uchio & F.Tokita : Antigenic reactivities of chemically modified β-lactoglobulins with antiserum to bovine β-lactoglobulin.
Agri. Biol. Chem. **49** 2531-2536 (1985)

[210] M.Bolognesi, J.Liberatori, R.Oberti & L.Ungaretti : Preliminary crystallographic data on buffalo β-lactoglobulin.
J. Mol. Biol. **131** 411-413 (1979)

[211] Unused.

[212] F.M.Young & A.C.T.North : Personal Communication (1987)

[213] M.Z.Papiz : Crystallographic studies of β-lactoglobulin at 2.8A resolution.
Ph.D. thesis, University of Edinburgh (1982)

[214] J.Godovac-Zimmermann, A.Conti, J.Liberatori & G.Braunitzer : Homology between the primary structures of β-lactoglobulins and human retinol-binding protein: evidence for a similar biological function?
Biol. Chem Hoppe-Seyler **366** 431-434 (1985)

[215] M.E.Newcomer, T.A.Jones, J.Aqvist, J.Sundelin, U.Eriksson, L.Rask & P.A.Peterson : The three-dimensional structure of retinol-binding protein.
EMBO J. **3** 1451-1454 (1984)

[216] J.Horwitz & J.Heller : Properties of the chromophore binding site of retinol-binding protein from human plasma.
J. Biol. Chem. **249** 4712-4719 (1974)

[217] M.A.Gawinowicz & D.S.Goodman : The site of linkage of a retinoid affinity label to plasma retinol-binding protein.
J. Prot. Chem. **4** 199-213 (1985)

[218] J.Aqvist, P.Sandblom, T.A.Jones, M.E.Newcomer, W.F.van Gunsteren & O.Tapia : Molecular dynamics simulations of the holo and apo forms of retinol binding protein. Structural and dynamical changes induced by retinol removal.
J. Mol. Biol. **192** 593-604 (1986)

[219] H.M.Holden, J.H.Law & I.Rayment : Crystallization of insecticyanin from the hemolymph of the tobacco hornworm Manduca sexta L. in a form

suitable for a high resolution structure determination.
J. Biol. Chem. **261** 4217-4218 (1986)

[220] K.Petratos, D.Tsernoglou & P.Cherbas : Preliminary characterization of crystals of the protein insecticyanin from the tobacco hornworm Manduca sexta L..
J. Mol. Biol. **189** 727 only (1986)

[221] R.Huber, M.Schneider, O.Epp, I.Mayr, A.Messerschmidt & J.Pflugrath : Crystallization, crystal structure analysis and preliminary molecular model of the bilin binding protein from the insect Pieris brassicae.
J. Mol. Biol. **195** 423-434 (1987)

[222] H.M.Holden, W.R.Rypniewski, J.H.Law & I.Rayment : The molecular structure of insecticyanin from the tobacco hornworm Manduca sexta L. at 2.6A resolution.
EMBO J. **6** 1565-1570 (1987)

[223] R.Huber, M.Schneider, I.Mayr, R.Muller, R.Deutzmann, F.Suter, H.Zuber, H.Falk & H.Kayser : Molecular structure of the bilin binding protein (BBP) from Pieris brassicae after refinement at 2.0A resolution.
J. Mol. Biol. **198** 499-513 (1987)

[224] M.Kumar, N.N.Dastur & M.V.Bhatt : Studies on the pigments of buffalo milk. I. Identification of biliverdin IXγ isomer in fresh buffalo milk.
Milchwissenschaft **39** 29-31 (1984)

[225a] D.E.McRee, J.A.Tainer, T.E.Meyer, J.van Beemen, M.A.Cusanovich & E.D.Getzoff : Crystallographic structure of a photoreceptor protein at 2.4A resolution.
Proc. Natl. Acad. Sci. **86** 6533-6537 (1989)

[225b] J.C.Sacchettini, J.I.Gordon & L.J.Banaszak : The structure of crystalline Escherichia coli-

derived rat intestinal fatty acid-binding protein at 2.5A resolution.
J. Biol. Chem. **263** 5815-5819 (1988)

[225c] T.A.Jones, T.Bergfors, J.Sedzik & T.Unge : The three-dimensional structure of P2 myelin protein.
EMBO J. **7** 1597-1604 (1988)

[226a] L.Rask, H.Anundi & P.A.Peterson : The primary structure of the human retinol-binding protein.
FEBS. Lett. **104** 55-58 (1979)

[226b] C.T.Riley, B.K.Barbeau, P.S.Keim, F.J.Kezdy, R.L.Heinrikson & J.H.Law : The covalent protein structure of insecticyanin, a blue biliprotein from the hemolymph of the tobacco hornworm, Manduca sexta L..
J. Biol. Chem. **259** 13159-13165 (1984)

[227] F.Sutur, H.Kayser & H.Zuber : The complete amino-acid sequence of the bilin-binding protein from Pieris brassicae and its similarity to a family of serum transport proteins like retinol-binding protein.
Biol. Chem. Hoppe-Seyler **369** 497-505 (1988)

[228] A.C.T.North : Three-dimensional arrangement of conserved amino acid residues in a superfamily of specific ligand-binding proteins.
Int. J. Biol. Macromol. **11** 56-58 (1989)

[229] J.C.Sacchettini, J.I.Gordon & L.J.Banaszak : Crystal structure of rat intestinal fatty-acid binding protein. Refinement and analysis of the Escherichia coli-derived protein with bound palmitate.
J. Mol. Biol. **208** 327-339 (1989)

[230] J.B.Findlay : Personal Communication (1987)

[231] D.R.Flower : Crystallographic studies of a family of homologous ligand binding proteins.
Ph.D. (first year report), University of Leeds (1989)

[232] P.Berman, P.Gray, E.Chen, K.Keyser, D.Ehrlich, H.Karten, M.Lacorbiere, F.Esch & D.Schubert : Sequence analysis, cellular localization, and expression of a neuroretina adhesion and cell survival molecule.
Cell **51** 135-142 (1987)

[233] T.Takagi, K.Takagi & T.Kawai : Complete amino acid sequence of human α_1-microglobulin.
Biochem. Biophys. Res. Comm. **98** 997-1001 (1981)

[234] C.Lopez, A.Grubb, F.Soriano & E.Mendez : The complete amino acid sequence of human complex-forming glycoprotein heterogeneous in charge (protein HC).
Biochem. Biophys. Res. Comm. **103** 919-925 (1981)

[235] J.-A.Haefliger, D.Jenne, K.K.Stanley & J.Tschopp : Structural homology of human complement component C8γ and plasma protein HC: identity of the cysteine bond pattern.
Biochem. Biophys. Res. Comm. **149** 750-754 (1987)

[236] F.Schoentgen, M.-H.Metz-Boutigue, J.Jolles, J.Constans & P.Jolles : Complete amino acid sequence of human vitamin D-binding protein (group-specific component): evidence of a three-fold internal homology as in serum albumin and α-fetoprotein.
Biochim. Biophys. Acta **871** 189-198 (1986)

[237] C.E.Argarana, I.D.Kuntz, S.Birken, R.Axel & C.R.Cantor : Molecular cloning and nucleotide sequence of the streptavidin gene.
Nucl. Acid Res. **14** 1871-1882 (1986)

[238] W.A.Hendrickson, A.Pahler, J.L.Smith, Y.Satow, E.A.Merritt & R.P.Phizackerley : Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation.
Proc. Natl. Acad. Sci. **86** 2190-2194 (1989)

[239] S.Pervaiz & K.Brew : Homology and structure-function correlations between α_1-acid glycoprotein and serum retinol-binding protein and its relatives.
Fed. Am. Soc. Exp. Biol. **1** 209-214 (1987)

[240] M.D.Perez, C.Diaz de Villegas, L.Sanchez, P.Aranda, J.M.Ena & M.Calvo : Interaction of fatty acids with β-lactoglobulin and albumin from ruminant milk.
J. Biochem. **106** 1094-1097 (1989)

[241] Unused.

[242] Unused.

[243] J.Pevsner, V.Hou, A.M.Snowman & S.N.Snyder : Odorant-binding protein.
J. Biol. Chem. **265** 6118-6125 (1990)

[244] M.C.Peitsch & M.S.Boguski : Is apolipoprotein D a mammalian bilin binding protein?
New Biologist **2** 1-10 (1990)

[245] S.Ali : Structure and expression of the gene encoding ovine β-lactoglobulin.
Ph.D. thesis, University of Edinburgh (1989)

[246] D.A.Sweetser, R.O.Heuckeroth & J.I.Gordon : The metabolic significance of mammalian fatty-acid-binding proteins.
Ann. Rev. Nutr. **7** 337-359 (1987)

[247] A.Cavaggioni, J.B.C.Findlay & R.Tirindelli : Personal Communication (1990)

[248] P.N.Shaw, W.A.Held & N.D.Hastie : The gene family major urinary proteins: expression in several secretory tissues of the mouse.
Cell **32** 755-761 (1983)

[249] P.E.Fielding & C.J.Fielding : A cholesteryl ester transfer complex in human plasma.
Proc. Natl. Acad. Sci. **77** 3327-3330 (1980)

[250] D.T.Drayna, J.W.McLean, K.L.Wion, J.M.Trent, H.A.Drabkin & R.M.Lawn : Human apolipoprotein D gene: gene sequence, chromosome localization, and homology to the α_{2u}-globulin superfamily.

DNA **6** 199-204 (1987)

[251] D.Drayna, C.Fielding, J.McLean, B.Baer, G.Casro, E.Chen,L.Comstock, W.Henzel, W.Kohr, L.Rhee, K.Wion & R.Lawn : Cloning and expression of human apolipoprotein D cDNA.
J. Biol. Chem. **261** 16535-16539 (1986)

[252] Unused.

[253] K.-H.Lee, R.G.Wells & R.R.Reed : Isolation of an olfactory cDNA: similarity to retinol-binding protein suggests a role in olfaction.
Science **235** 1053-1056 (1987)

[254] J.Pevsner, R.R.Trifiletti, S.M.Strittmatter & S.H.Snyder : Isolation and characterization of an olfactory receptor protein for odorant pyrazines.
Proc. Natl. Acad. Sci. **82** 3050-3054 (1985)

[255] E.Bignetti, A.Cavaggioni, P.Pelosi, K.C.Persaud, R.T.Sorbi & R.Tirindelli : Purification and characterisation of an odorant-binding protein from cow nasal tissue.
Eur. J. Biochem. **149** 227-231 (1985)

[256] L.J.Haars & H.C.Pitot : Hormonal and developmental regulation of glycosylated α_{2u}-globulin synthesis.
Arch. Biochem. Biophys. **201** 556-563 (1980)

[257] J.Pevsner, R.R.Reed, P.G.Feinstein & S.H.Synder : Molecular cloning of odorant-binding protein: member of a ligand carrier family.
Science **241** 336-339 (1988)

[258] A.Cavaggioni, R.T.Sorbi, J.N.Keen, D.J.C.Pappin & J.B.C.Findlay : Homology between the pyrazine-binding protein from nasal mucosa and major urinary proteins.
FEBS Lett. **212** 225-228 (1987)

[259] G.Vandoren, B.Mertens, W.Heyns, H.van Baelen, W.Rombauts & G.Verhoeven : Different forms of α_{2u}-globulin in male and female rat urine.
Eur. J. Biochem. **134** 175-181 (1983)

[260] R.D.Unterman, K.R.Lynch, H.L.Nakhasi, K.P.Dolan, J.W.Hamilton, D.V.Cohn & P.Feigelson : Cloning and sequence of several α_{2u}-globulin cDNAs.
Proc. Natl. Acad. Sci. **78** 3478-3482 (1981)

[261] E.A.Lock, M.Charbonneau, J.Strasser, J.A.Swenberg & J.S.Bus : 2,2,4-trimethylpentane-induced nephrotoxicity. II. The reversible binding of a TMP metabolite to a renal protein fraction containing α_{2u}-globulin.
Toxicol. Appl. Pharm. **91** 182-192 (1987)

[262] D.E.Brooks, A.R.Means, E.J.Wright, S.P.Singh & K.K.Tiver : Molecular cloning of the cDNA for two major androgen-dependent secretory proteins of 18.5 kilodaltons synthesized by the rat epididymis.
J. Biol. Chem. **261** 4956-4961 (1986)

[263] D.E.Brooks : The major androgen-regulated secretory proteins of the rat epididymis bear sequence homology with members of the α_{2u}-globulin superfamily.
Biochem. Int. **14** 235-240 (1987)

[264] W.J.Henzel, H.Rodriguez, A.G.Singer, J.T.Stults, F.Macrides, W.C.Agosta & H.Naill : The primary structure of aphrodisin.
J. Biol. Chem. **263** 16682-16687 (1988)

[265] Unused.

[266] S.C.Bell & S.R.Patel : Immunochemical detection, physico-chemical characterization and levels of pregnancy-associated endometrial α_2-globulin (α_2-PEG) in seminal plasma of men.
J. Reprod. Fert. **80** 31-42 (1987)

[267] M.L.Friedman, K.T.Schlueter, T.L.Kirley & H.B.Halsll : Fluorescence quenching of human orosomucoid. Accessibility to drugs and small quenching agents.
Biochem. J. **232** 863-867 (1985)

[268] M.Ganguly, R.H.Carnighan & U.Westphal : Steroid-protein interactions. XIV. Interaction between human α_1-acid glycoprotein and progesterone.
Biochemistry **6** 2803-2814 (1967)

[269] Unused.

[270] R.Cooper, D.M.Eckley & J.Papaconstantinou : Nucleotide sequence of the mouse α_1-acid glycoprotein gene 1.
Biochemistry **26** 5244-5250 (1987)

[271] L.J.Alexander, G.Hayes, M.J.Pearse, C.W.Beattie, A.F.Stewart, I.M.Willis & A.G.Mackinlay : Complete sequence of the bovine β-lactoglobulin cDNA.
Nucl. Acid Res. **17** 6739 only (1989)

[272] A.C.T.North : Applications of molecular graphics for the study of recognition.
J. Mol. Graphics **7** 67-70 (1989)

[273] J.-C.Mercier, P.Gaye, S.Soulier, D.Hue-Delahaie & J.-L.Vilotte : Construction and identification of recombinant plasmids carrying cDNAs coding for ovine α_{s1}-, α_{s2}-, β-, κ-casein and β-lactoglobulin. Nucleotide sequence of α_{s1}-casein cDNA.
Biochimie **67** 959-971 (1985)

[274] P.Gaye, D.Hue-Delahaie, J.-C.Mercier, S.Soulier, J.-L.Vilotte & J.P.Furet : Ovine β-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland.
Biochimie **68** 1097-1107 (1986)

[275] S.Ali & A.J.Clark : Characterization of the gene encoding ovine beta-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins.
J. Mol. Biol. **199** 415-426 (1988)

[276] I.M.Willis, A.F.Stewart, A.Caputo, A.R.Thompson & A.G.Mackinlay : Construction and identification by partial nucleotide sequence analysis of bovine

casein and β-lactoglobulin cDNA clones.
DNA **1** 375-386 (1982)

[277] V.N.Ivanov, E.S.Judinkova, S.I.Gorodetsky : Molecular cloning of bovine β-lactoglobulin cDNA.
Biol. Chem. Hoppe-Seyler **369** 425-429 (1988)

[278] A.C.Jamieson, M.A.Vandeyar, Y.C.Kang, J.E.Kinsella & C.A.Batt : Cloning and nucleotide sequence of the bovine β-lactoglobulin gene.
Gene **61** 85-90 (1987)

[279] H.M.Said, D.E.Ong & J.L.Shingleton : Intestinal uptake of retinol: enhancement by bovine milk β-lactoglobulin.
Am. J. Clin. Nutr. **49** 690-694 (1989)

[280] A.J.Wonacott : Geometry of the rotation method.
The rotation method in crystallography **Chap 7** 75-103 (1977)

[281] J.R.Helliwell : Synchrotron x-radiation protein crystallography.
Rep. Prog. Phys. **47** 1403-1497 (1984)

[282] Unused.

[283] S.G.Hambling, S.J.Yewdall, A.C.T.North & L.Sawyer : Comparison of two crystal forms of bovine β-lactoglobulin.
Acta Cryst. **A43** (Supplement) C-26 only (1987)

[284] Unused.

[285] A.G.W.Leslie : Profile-fitting.
Proc. Daresbury Study Weekend 39-50 (1987)

[286] M.G.Rossmann : Processing oscillation diffraction data for very large unit cells with an automatic convolution technique and profile fitting.
J. Appl. Cryst. **12** 225-238 (1979)

[287] Unused.

[288] Unused.

[289] B.W.Matthews, C.E.Klopfenstein & P.M.Colman : A computer controlled film scanner for x-ray crystallography.

J. Phys. E: Sci. Instru. **5** 353-359 (1972)

[290] R.Khan, R.Fourme, A.Gadet, J.Janin, C.Dumas & D.Andre : Macromolecular crystallography with synchrotron radiation: photographic data collection and polarization correction.
J. Appl. Cryst. **15** 330-337 (1982)

[291] Unused.

[292] G.C.Fox & K.C.Holmes : An alternative method of solving the layer scaling equations of Hamilton, Rollett and Sparks.
Acta Cryst. **20** 886-891 (1966)

[293] Unused.

[294] S.Neidle : On the determination of heavy-atom positions in various elastase derivatives.
Acta Cryst. **B29** 2645-2647 (1973)

[295] E.J.Dodson : Molecular replacement: the method and its problems.
Proc. Daresbury Study Weekend 33-45 (1985)

[296] J.C.Phillips & K.O.Hodgson : The use of anomalous scattering effects to phase diffraction patterns from macromolecules.
Acta Cryst. **A36** 856-864 (1980)

[297] U.Cogan, M.Kopelman, S.Mokady & M.Shinitzky : Binding affinities of retinol and related compounds to retinol binding proteins.
Eur. J. Biochem. **65** 71-78 (1976)

[298] G.A.Sim : The distribution of phase angles for structures containing heavy atoms. II. Modification of the normal heavy atom method for non-centro-symmetrical structures.
Acta Cryst. **12** 813-815 (1959)

[299] W.A.Hendrickson & E.E.Lattman : Representation of phase probability distributions for simplified combination of independent phase information.
Acta Cryst **B26** 136-143 (1970)

[300] D.Stuart & P.Artymiuk : The use of phase combination

in crystallographic refinement: the choice of amplitude coefficients in combined syntheses.
Acta Cryst. **A40** 713-716 (1985)

[301] T.A.Jones : A graphics model building and refinement system for macromolecules.
J. Appl. Cryst. **11** 268-272 (1978)

[302] R.Diamond : A real space refinement procedure for proteins.
Acta Cryst. **A27** 436-452 (1971)

[303] A.D.Podjarny, T.N.Bhat & M.Zwick : Improving crystallographic macromolecular images: the real-space approach.
Ann. Rev. Biophys. Biophys. Chem. **16** 351-373 (1987)

[304] E.M.C.Turner : Crystallographic studies of β-lacto-globulin modified by carboxypeptidase.
Ph.D. thesis, University of Edinburgh (1975)

[305] J.H.Konnert & W.A.Hendrickson : A restrained-parameter thermal-factor refinement procedure.
Acta Cryst **A36** 344-350 (1980)

[306] D.E.Tronrud, L.F.Ten Eyck & B.W.Matthews : An efficient general-purpose least-squares refinement program for macromolecule structures.
Acta Cryst. **A43** 489-501 (1987)

[307] R.C.Agarwal : A new least-squares refinement technique based on the fast Fourier transform algorithm.
Acta Cryst **A34** 791-809 (1978)

[308] P.Gros, M.Fujinaga, A.Mattevi, F.M.D.Vellieux, W.F.van Gunsteren & W.G.J.Hol : Protein structure refinement by molecular dynamics techniques.
Proc. Daresbury Study Weekend 1-15 (1989)

[309] A.T.Brunger, J.Kuriyan & M.Karplus : Crystallographic R factor refinement by molecular dynamics.
Science **235** 458-460 (1987)

[310] W.G.J.Hol, L.M.Halie & C.Saunder : Dipoles of the α-helix and β-sheet: their role in protein folding.

Nature **294** 532-536 (1981)

[311] A.C.Ghose, S.Chaudhuri & A.Sen : Hydrogen ion equilibrium and sedimentation behaviour of goat β-lactoglobulin.
Arch. Biochem. Biophys. **126** 232-243 (1968)

[312] P.Bayley & S.Martin : Personal Communication (1986)

[313] P.M.Bayley : The analysis of circular dichroism of biomolecules.
Prog. Biophys. Mol. Biol. **27** 1-76 (1973)

[314] H.A.McKenzie : Personal Communication (1988)

[315] Unused.

[316] Unused.

[317] F.W.J.Teale & G.Weber : Ultraviolet fluorescence of the aromatic amino acids.
Biochem. J. **65** 476-482 (1957)

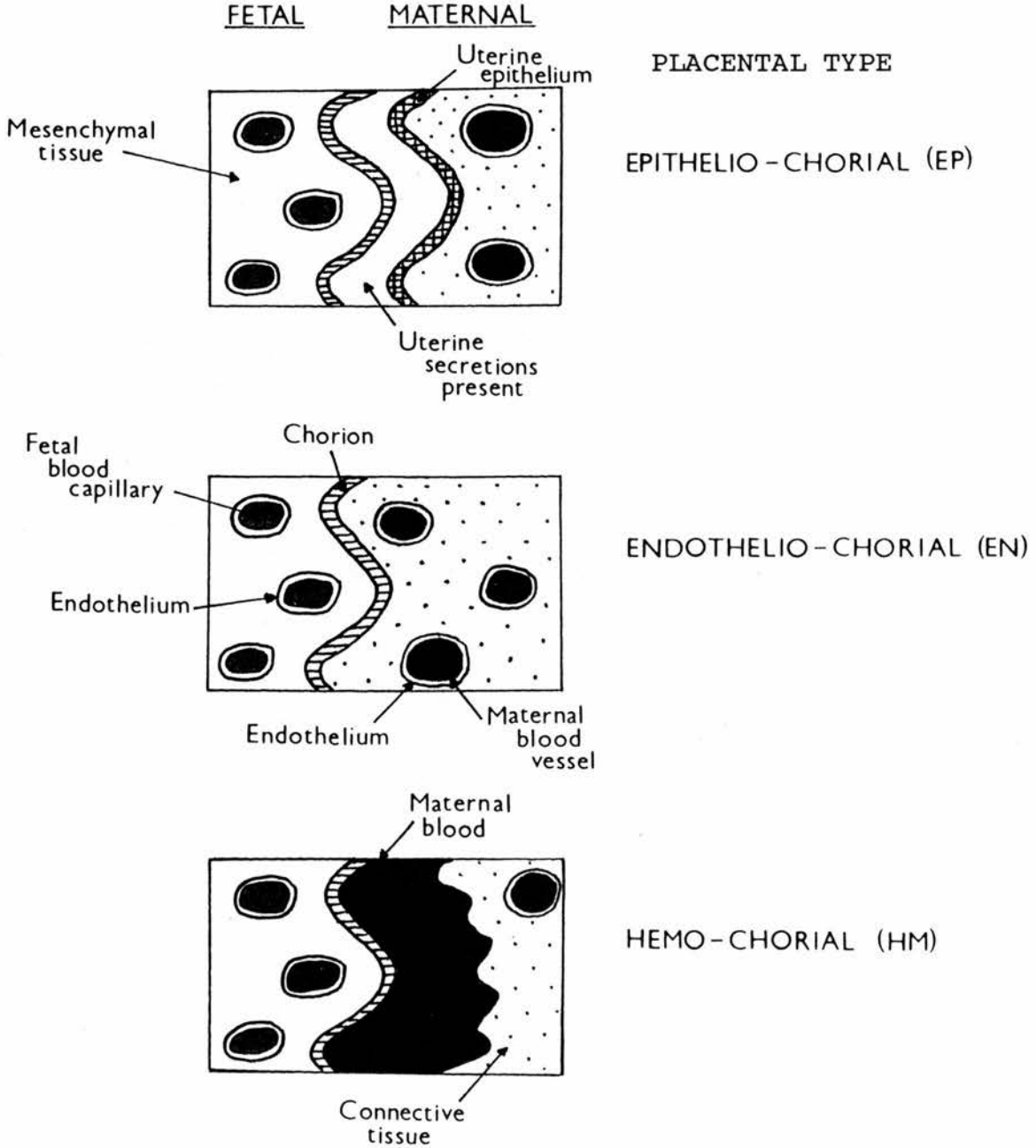
[318] M.R.Eftink & C.A.Ghiron : Fluorescence quenching studies with proteins.
Anal. Biochem. **114** 199-227 (1981)

[319] G.Allen : Sequencing of proteins and peptides.
Laboratory Techniques in Biochemistry and Molecular Biology **Vol 9** Chap 6 (Elsevier 1989)

APPENDIX 1

SOME PLACENTAL TYPES OF MAMMALS.

The placental type of a mammal depends upon the intimacy of contact between the fetal and maternal circulations. There can be upto six layers of tissue separating the maternal and fetal blood: the endothelium of the mother's blood vessels, connective tissue, the uterine epithelium, the chorion, mesenchymal tissue, and the endothelium of the fetal blood capillaries. The progressive removal of these layers is illustrated below.



APPENDIX 2

SIMPLIFIED THEORY OF X-RAY DIFFRACTION.

When matter is irradiated with x-rays the electrons surrounding the atoms are made to oscillate and emit a secondary ray. The combined contributions of all these elastically scattering electrons produce a x-ray diffraction pattern.

If an incident x-ray beam strikes a volume element dr, containing matter of electron density $\rho(\underline{r})$, then the contribution of this volume element to the diffraction pattern is:

$$\rho(\underline{r})\exp\ 2\pi i\underline{s}.\underline{r}\ dr$$

where $\underline{s}$ is the scattering vector and $\underline{r}$ is the position vector of the volume element relative to a common origin. The scalar quantity $\underline{s}.\underline{r}$ determines the phase of each volume element's contribution to the diffraction pattern.

The integrated contribution of all matter in volume V is given by

$$\underline{G}(s) = \int \rho(\underline{r})\ \exp\ 2\pi i\underline{s}.\underline{r}\ dr$$

The diffracting material of a crystal is the matter of the unit cell convoluted with the crystalline lattice. If $\underline{H}(s)$ is the Fourier transform of the function that generates the crystal lattice along which the unit cell can be translated to produce the crystal volume, then it can be shown that the crystalline Fourier transform $\underline{F}(s)$ is given by

$$\underline{F}(s) = \underline{G}(s)\underline{H}(s)$$

The integral Miller indices (hkl) of the reciprocal lattice points are given by the Laue conditions for diffraction

$$\underline{s}.\underline{a} = h;\ \underline{s}.\underline{b} = k;\ \underline{s}.\underline{c} = l$$

where $\underline{a}$, $\underline{b}$, and $\underline{c}$ are the real cell unit vectors. If these lattice points are given an arbitary weight then $\underline{H}(hkl) = 1$ and

$$\underline{F}(hkl) = \int \rho(\underline{r}) \exp 2\pi i(hx_j + ky_j + lz_j)$$

where x_j, y_j and z_j are the fractional cell coordinates. This integration can be reduced to a summation, and hence the electron density at any point can be given by

$$\rho(x_j,y_j,z_j) = (1/V) \Sigma \underline{F}(hkl) \exp -2\pi i(hx_j + ky_j + lz_j)$$

In general $\underline{F}(hkl)$ is a complex number and can be represented by

$$\underline{F}(hkl) = F(hkl) \exp 2\pi i\alpha(hkl)$$

where $\alpha(hkl)$ and $F(hkl)$ are the phase and amplitude of the structure factor respectively.

The problem that faces all x-ray diffraction studies is how to obtain estimates of the phases. This difficulty arises because only the intensity of a diffraction pattern (which is equal to the square of the structure factor amplitude) can be measured.

APPENDIX 3

PROGRAM, MERGE, FOR REJECTING INCONSISTENT REFLECTIONS.

```
C
C CRITERIA FOR REJECTION ARE SIMILAR TO THOSE IN AGROVATA
C
      DIMENSION F(4),S(4),N(4),IDATA(12),LOOKUP(12),
     1  CELL(6),X(4),ISET(4),ITOT(4)
      CHARACTER*4 LOOK(11),TITLE(18)
      DATA  LOOK/'H  K  ','L  S  ','F  SI  ','GF  F',
     1  'C  SI','GFC ','FR  S','IGFR','  FD  ','SIGF',
     1  'D    '/,ISET/4*0/,ITOT/4*0/
C
      READ(5,1000) TITLE
1000  FORMAT(18A4)
      CALL SRLCF1(11,'HKLIN',44,LOOK,LOOKUP,.TRUE.,
     1  NCOLS,CELL)
      CALL HDLCF1(12,'HKLOUT',72,TITLE,44,LOOK,CELL,0,1)
C
100   CALL RDLCF1(IDATA,*800,*200)
      I=0
      FSUM=0.
      L=0
      DO 101 J=5,11,2
      I=I+1
      S(I)=2.0*IDATA(J)*IDATA(J+1) + IDATA(J+1)**2
      F(I)=IDATA(J)**2
      FSUM=FSUM+F(I)
      IF (IDATA(J+1).NE.0) THEN
      L=L+1
      ITOT(I)=ITOT(I)+1
      ENDIF
101   CONTINUE
      IF (L.EQ.0) GO TO 100
      IF (L.EQ.1) THEN
102   CALL WLCF1 (IDATA)
      GO TO 100
      ENDIF
C DETERMINE THE REJECTION PARAMETER DEL
      FSUM=FSUM/FLOAT(L)
      K=0
      DO 105 J=1,4
      X(J)=0.
      IF (S(J).EQ.0) GO TO 105
      X(J)=(FSUM-F(J))/S(J)
      IF (ABS(X(J)).LE.3) K=K+1
105   CONTINUE
      IF (K.EQ.L) GO TO 102
C REJECTION REQUIRED, WHICH?!
      IF (L.EQ.2) THEN
C 2 REFLECTIONS, 1 OR BOTH FAIL REJECTION CRITERIA
      IF (K.EQ.1) THEN
      DO 107, I=1,4
```

```
        IF (X(I).GT.3.) THEN
        WRITE (6,106) IDATA
106     FORMAT (16I5)
        IDATA(2*(I-1)+5)=0
        IDATA(2*(I-1)+6)=0
        ISET(I)=ISET(I)+1
        ENDIF
107     CONTINUE
        GO TO 102
        ELSE
C  BOTH FAIL
        Y=99999999.
        DO 108 I=1,4
        IF (F(I).LT.Y.AND.S(I).NE.0) J=I
108     CONTINUE
        WRITE (6,109)IDATA
109     FORMAT(16I5)
        IDATA(2*(J-1)+5)=0
        IDATA(2*(J-1)+6)=0
        ISET(J)=ISET(J)+1
        ENDIF
        GO TO 102
        ELSE
C   3 OR 4 REFLECTIONS
        Y=0.
        DO 110 I=1,4
        IF (X(I).GT.Y) THEN
        Y=X(I)
        J=I
        ENDIF
110     CONTINUE
        WRITE (6,111)IDATA
111     FORMAT(16I5)
        IDATA(2*(J-1)+5)=0
        IDATA(2*(J-1)+6)=0
        ISET(J)=ISET(J)+1
        GO TO 102
        ENDIF
200     CONTINUE
        CALL CRLCF1
        CALL CWLCF1
        WRITE (6,206)ISET
206     FORMAT(' NUMBER JUNKED FOR EACH DATASET',4I6)
        WRITE (6,207)ITOT
207     FORMAT(' OUT OF A TOTAL NUMBER OF ',4I6)
        STOP
800     WRITE (6,801)
801     FORMAT('  ** ERROR IN READING LCF FILE **')
        STOP
        END
```

APPENDIX 4

PROGRAM, MEANF, FOR AVERAGING THE STRUCTURE FACTORS FOR EACH REFLECTION.

```
C
      DIMENSION F(4),S(4),N(4),IDATA(12),LOOKUP(12),
     1  CELL(6),X(4),JDATA(6),W(4),WF(4),SIGF(4)
      CHARACTER*4 LOOK(11),TITLE(18)
      DATA  LOOK/'H  K  ','L  S  ','F  SI  ','GF  F',
     1  'C  SI','GFC ','FR  S','IGFR','  FD  ','SIGF',
     1  'D    '/
C
      READ(5,1000) TITLE
1000  FORMAT(18A4)
      CALL SRLCF1(11,'HKLIN',44,LOOK,LOOKUP,.TRUE.,
     1  NCOLS,CELL)
      CALL HDLCF1(12,'HKLOUT',72,TITLE,14,LOOK,CELL,0,1)
C
      K=0.
100   CALL RDLCF1(IDATA,*800,*200)
      WRITE (6,102)IDATA
102   FORMAT(16I5)
      K=K+1
      I=0.
      WSUM=0.
      WFSUM=0.
101   DO 103 J=5,11,2
      I=I+1
      S(I)=2.0*IDATA(J)*IDATA(J+1) + IDATA(J+1)**2
      IF (S(I).EQ.0) GO TO 103
      F(I)=IDATA(J)**2
C CALCULATE WEIGHTED MEAN F AND CORRESPONDING SIGF
      W(I)=1./(S(I)*S(I))
      WSUM=WSUM+W(I)
      WF(I)=W(I)*F(I)
      WFSUM=WFSUM+WF(I)
103   CONTINUE
      WGHTF=WFSUM/WSUM
      WGHTSF=SQRT(1./WSUM)
C CONVERT FROM INTENSITIES BACK TO F'S
      FBAR=SQRT(WGHTF)
      SFBAR= -FBAR+SQRT(WGHTF+WGHTSF)
C PREPARING TO WRITE LCF FILE
      DO 105 I=1,4
      JDATA(I)=IDATA(I)
105   CONTINUE
      JDATA(5)=FBAR
      JDATA(6)=SFBAR
102   CALL WLCF1 (JDATA)
      GO TO 100
200   CONTINUE
      CALL CRLCF1
```

```
      CALL CWLCF1
      STOP
800   WRITE (6,801)
801   FORMAT('  ** ERROR IN READING LCF FILE **')
      STOP
      END
```

APPENDIX 5

SUMMARY OF THE DETERMINATION OF THE LATTICE X STRUCTURE OF BOVINE BLG AT pH 6.5 [141].

Triclinic crystals of lattice X (space group P1: a=37.8A, b=49.6A, c=56.6A, α=123.4°, β=97.3°, γ=103.7°) were grown from a 4M ammonium sulphate solution at pH 6.5, and were typically of dimensions 0.7mm x 0.4mm x 0.3mm.

Data were collected to 2.0A resolution, from thirty-three crystals, on an Enraf-Nonius CAD4 diffractometer with copper Kα radiation (λ=1.5418A). Empirical absorption corrections were applied using absorption curves determined from the measurements of five axial reflections at different Bragg angles. The data were then processed using the MRC/CCP4 program suite - the integrated intensities of the reflections were obtained by profile-fitting, a correction for the Lorentz-polarization factor was applied, and then the batches scaled together using the Fox and Holmes scaling routine [292].

Number of measurements = 51283
Number of reflections = 21471
% of unique reflections = 98%
% > 3SD = 87.5%
R_{merge} (on I) = 0.058

The low resolution 6A structure [5] was extended to 3A by the inclusion of phases determined from a single isomorphous derivative. A rotation function was applied, to maximize the fitting of the lattice Y dimer into the lattice X triclinic cell. Each monomer of lattice X was then refitted separately using model-building and cycles of Hendrickson-Konnert refinement, and this reduced the R_F value from 0.57 to 0.41.

The coordinates of this initial model, a backbone

chain of 136 polyalanine residues per monomer, were then used to calculate combined phases. These were then concatenated to data from the new 2.0A dataset. Manual model building using the molecular graphics program FRODO, to replace the polyalanine residues by the correct amino acids; and iterative cycles of Hendrickson-Konnert LS refinement, were then applied. $3F_{obs}-2F_{calc}$ maps were generated at various stages and as refinement proceeded to 2.5A resolution the R_F dropped to 0.295. The inclusion of all the amplitudes to 2.0A resolution, and three cycles of LS refinement, has produced a current model with an R_F of 0.28, although the R_F of the highest resolution shell still remains high (0.44).